

影像食光-今晚吃甚麼

Visual Feast - What to Eat Tonight?

BDSE32-Group Two

組員介紹

Leader

機器學習工程



魯又齊

Executive

特徵工程

機器學習

深度學習

資料工程



王柏歲

Executive

網路爬蟲

資料清整

資料分析

系統工程



黃湘玲

Executive

系統工程

Hadoop架設

資料管理

組員介紹

資料工程



郭安迪

Executive

效能檢測

資料分析

後端工程



林聖尉

Executive

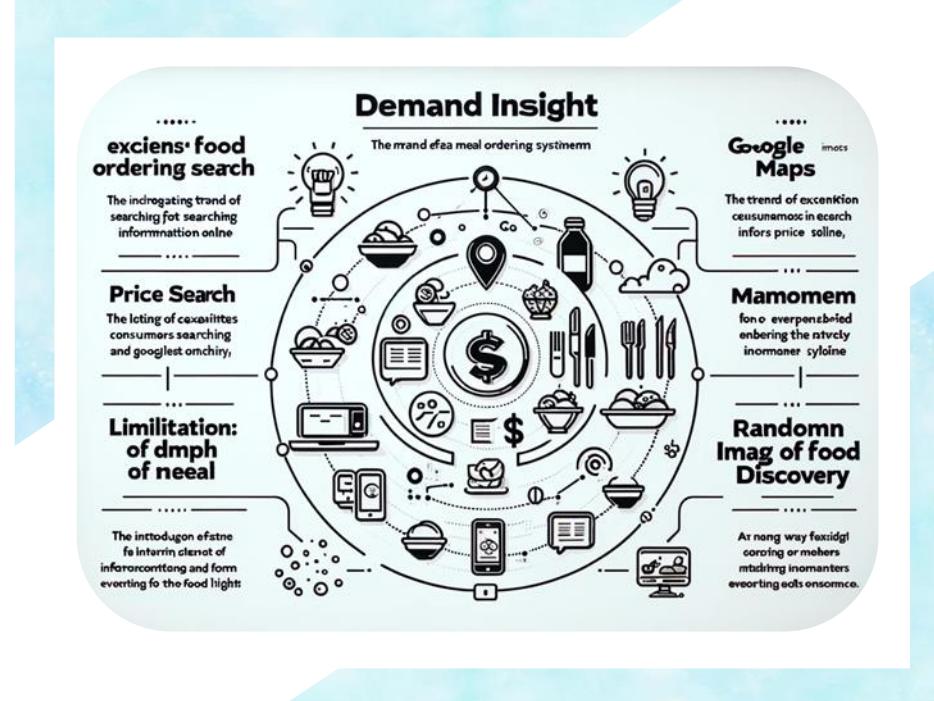
機器學習

資料處理

後端工程

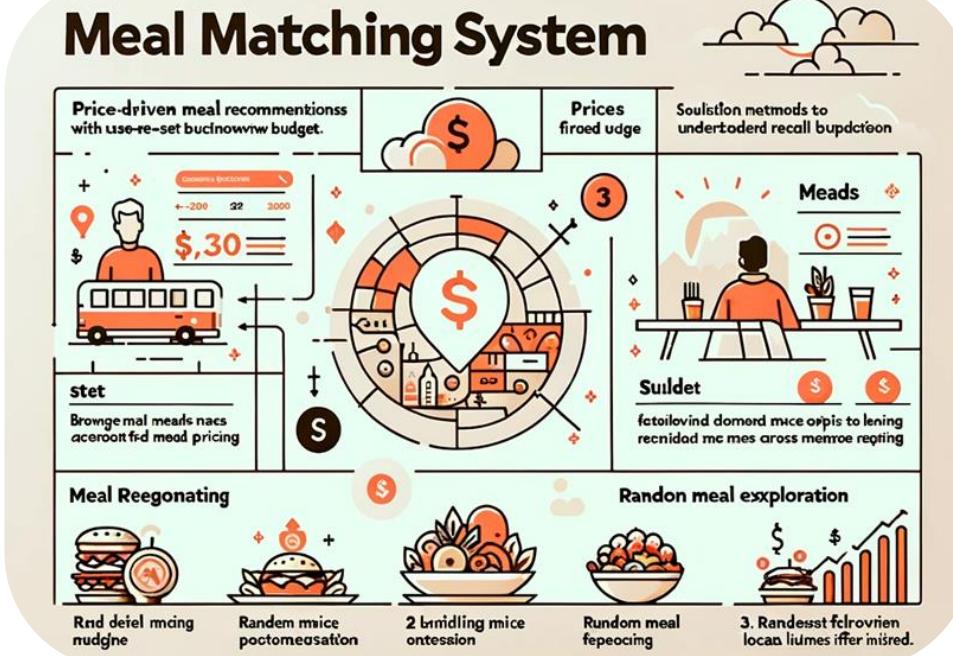
需求洞察

Insights



- 消費者透過網路尋找餐點信息的趨勢增加。
- 現有訂餐軟體和Google Maps在價格搜索和隨機餐點發現方面的限制。
- 「影像食光」提供新的餐點探索方式。

解決方案



Solution

- 1) 價格驅動的餐點推薦：**
 - ✓ 透過設定預算獲得餐點推薦。
 - ✓ 與傳統以餐廳為中心的推薦方式不同。
- 2) 餐點價格預測：**
 - ✓ 幫助消費者了解餐點的合理價格。
 - ✓ 為餐廳提供展示高價值餐點的洞察。
- 3) 隨機餐點探索：**
 - ✓ 隨機瀏覽全台的餐點。
 - ✓ 對感興趣的餐點，系統會尋找附近類似選擇。

市場定位

Positioning

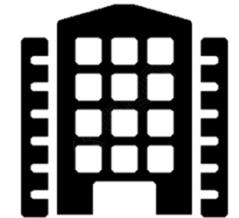


消費者

- 提供基於價格和餐點的個性化推薦。
- 滿足消費者探索新餐飲選擇的需求。

- 為餐廳業主提供市場洞察。
- 透過視覺化預測價格增加產品價值。

企業方



Scraping

網路爬蟲

Cleaning

資料清整



王柏威
Speaker

資料收集-使用工具



Selenium

BeautifulSoup



Requests
http for humans



pandas



pillow

資料來源 / 數據量



- ✓ 總計店家數：**41282家**
- ✓ 實際使用店家數：**28783家**
- ✓ 總計爬取餐點數：約**119萬筆**
- ✓ 實際使用餐點數：約**110萬筆**
- ✓ 總計留言數：約**363萬筆**



小吉咖哩 (台北八德店)

★ 4.7 (4000+)

\$\$ · 台式



讚: 美食 pasta!

\$\$ · 歐美 · <店內便



泰興 南洋料理

★ 4.4 (45)

\$\$\$ · 東南亞



S&D 咖啡 (松山)

\$\$ · 歐美

店內搜尋 人氣精選 (6) 咖哩 (13) 小菜 (5) 炸物 (7) 飲品/湯品 (6) 甜品 (3)

🔥 人氣精選 🔥

大家都點這些。金賀呷！手刀點起來

鹽酥雞咖哩

雞腿肉



從 \$ 162 \$ 170

豬排咖哩



從 \$ 171 \$ 180

土魠魚咖哩



從 \$ 171 \$ 180

小吉咖哩



從 \$ 133 \$ 140

共爬到41282家店

第一次蒐集

Code	名稱	網址	評分	餐點類型
w5x6	醬子義大利麵(新竹店)	/restaurant/w5x6/jiang-zi-yi-da-li-mian-xin-zhu-dian	4.5 stars out of 5	歐美
y94p	degugu 果果迪(新竹金山店)	/restaurant/y94p/degugu-guo-guo-di-xin-zhu-jin-shan	4.9 stars out of 5	飲料
ecot	TINHA MUSLIM RESTAURANT	/restaurant/ecot/tinha-muslim-restaurant	4.5 stars out of 5	東南亞
hrbu	甘蔗の媽媽(新竹光復店)	/restaurant/hrbu/gan-zhe-noma-ma-xin-zhu-guang-fu	4.9 stars out of 5	飲料
jvti	東家麵食	/restaurant/jvti/dong-jia-mian-shi	4.9 stars out of 5	台式
enrc	陳飲CHECKIN(新竹金山店)	/restaurant/enrc/chen-yin-checkin-xin-zhu-jin-shan-di	4.8 stars out of 5	飲料
j5zu	軟蛋醬 蛋包飯專門製作所	/restaurant/j5zu/ruan-dan-jiang-dan-bao-fan-zhuan-m	4.8 stars out of 5	台式
fmt4	珍煮丹(新竹光復店)	/restaurant/fmt4/zhen-zhu-dan-xin-zhu-guang-fu-dian	4.8 stars out of 5	飲料
n724	釜山崔炸雞(新竹金山店)	/restaurant/n724/fu-shan-cui-zha-ji-xin-zhu-jin-shan-q	4.7 stars out of 5	韓式
xqya	蟬吃茶天然茶飲(新竹竹科店)	/restaurant/xqya/chan-chi-cha-tian-ran-cha-yin-xin-zh	4.9 stars out of 5	飲料
u9e0	五桐號 WooTEA(竹金山旗艦店)	/restaurant/u9e0/wu-tong-hao-wootea-zhu-jin-shan-q	4.9 stars out of 5	飲料
v1bv	初蛋糕・豆漿蛋糕達人(新竹關東店)	/restaurant/v1bv/chu-dan-gao-dou-jiang-dan-gao-da-r	4.8 stars out of 5	甜點
a6ab	Super.tea	/restaurant/a6ab/super-tea	4.7 stars out of 5	飲料

- 從各縣市主頁蒐集所有店家

網路爬蟲

Food Panda

共爬到119萬筆餐點

第二次蒐集

pic_id	food_name	food_id	introduce	price	address	latitude	longitude	label	shop_id
h2jy_16820023.jpg	雞肉飯	16820023	豬脂（油）原	36	(△)臺南市安南區安和路一段	23.03077	120.2114	Taiwanese,Bento	h2jy
h2jy_16820024.jpg	魚鬆肉燥飯	16820024	肉品原產地：	31.5	(△)臺南市安南區安和路一段	23.03077	120.2114	Taiwanese,Bento	h2jy
h2jy_16820025.jpg	貢丸湯	16820025	肉品原產地：	31.5	(△)臺南市安南區安和路一段	23.03077	120.2114	Taiwanese,Bento	h2jy
h2jy_16820026.jpg	魚丸湯	16820026	豬脂（油）原	31.5	(△)臺南市安南區安和路一段	23.03077	120.2114	Taiwanese,Bento	h2jy
h2jy_16820027.jpg	味噌湯	16820027	溫暖且富含蛋	18	(△)臺南市安南區安和路一段	23.03077	120.2114	Taiwanese,Bento	h2jy
h2jy_16820028.jpg	筍干	16820028	豬脂（油）原	18	(△)臺南市安南區安和路一段	23.03077	120.2114	Taiwanese,Bento	h2jy
h2jy_16820029.jpg	滷蛋	16820029	豬脂（油）原	15.3	(△)臺南市安南區安和路一段	23.03077	120.2114	Taiwanese,Bento	h2jy
h2jy_16820030.jpg	竹輪	16820030	豬脂（油）原	13.5	(△)臺南市安南區安和路一段	23.03077	120.2114	Taiwanese,Bento	h2jy
h2jy_16820031.jpg	滷丸	16820031	豬脂（油）原	9	(△)臺南市安南區安和路一段	23.03077	120.2114	Taiwanese,Bento	h2jy
h2jy_16820032.jpg	油豆腐	16820032	豬脂（油）原	7.2	(△)臺南市安南區安和路一段	23.03077	120.2114	Taiwanese,Bento	h2jy
w60j_54826892.jpg	夏威夷披薩	54826892	11吋 整片300	297	(△)臺南市中西區健康路一段	22.98147	120.1978	Pizza,American	w60j
w60j_54826893.jpg	BBQ雞披薩	54826893	11吋 整片300	297	(△)臺南市中西區健康路一段	22.98147	120.1978	Pizza,American	w60j
w60j_54826895.jpg	德式香腸披薩	54826895	11吋 整片300	297	(△)臺南市中西區健康路一段	22.98147	120.1978	Pizza,American	w60j
w60j_54826896.jpg	起司起司披薩	54826896	11吋 整片300	297	(△)臺南市中西區健康路一段	22.98147	120.1978	Pizza,American	w60j
w60j_54826894.jpg	哈燒辣雞披薩	54826894	11吋 整片300	297	(△)臺南市中西區健康路一段	22.98147	120.1978	Pizza,American	w60j

- 使用第一表單的Code · 透過Food Panda網址API · 發送Request

資料清整

FoodPanda

清整前

pic_id	food_name	food_id	introduce	price	address	latitude	longitude	label	shop_id
h2jy_16820023.jpg	雞肉飯	16820023	豬脂（油）原	36	(△)臺南市安南區安和路一段	23.03077	120.2114	Taiwanese,Bento	h2jy
h2jy_16820024.jpg	魚鬆肉燥飯	16820024	肉品原產地：	31.5	(△)臺南市安南區安和路一段	23.03077	120.2114	Taiwanese,Bento	h2jy
h2jy_16820025.jpg	貢丸湯	16820025	肉品原產地：	31.5	(△)臺南市安南區安和路一段	23.03077	120.2114	Taiwanese,Bento	h2jy
h2jy_16820026.jpg	魚丸湯	16820026	豬脂（油）原	31.5	(△)臺南市安南區安和路一段	23.03077	120.2114	Taiwanese,Bento	h2jy
h2jy_16820027.jpg	味噌湯	16820027	溫暖且富含蛋	18	(△)臺南市安南區安和路一段	23.03077	120.2114	Taiwanese,Bento	h2jy
h2jy_16820028.jpg	筍干	16820028	豬脂（油）原	18	(△)臺南市安南區安和路一段	23.03077	120.2114	Taiwanese,Bento	h2jy
h2jy_16820029.jpg	滷蛋	16820029	豬脂（油）原	15.3	(△)臺南市安南區安和路一段	23.03077	120.2114	Taiwanese,Bento	h2jy
h2jy_16820030.jpg	竹輪	16820030	豬脂（油）原	13.5	(△)臺南市安南區安和路一段	23.03077	120.2114	Taiwanese,Bento	h2jy
h2jy_16820031.jpg	滷丸	16820031	豬脂（油）原	9	(△)臺南市安南區安和路一段	23.03077	120.2114	Taiwanese,Bento	h2jy
h2jy_16820032.jpg	油豆腐	16820032	豬脂（油）原	7.2	(△)臺南市安南區安和路一段	23.03077	120.2114	Taiwanese,Bento	h2jy
w60j_54826892.jpg	夏威夷披薩	54826892	11吋 整片30CM大	297	(△)臺南市中西區健康路一段	22.98147	120.1978	Pizza,American	w60j
w60j_54826893.jpg	BBQ雞披薩	54826893	11吋 整片30CM大	297	(△)臺南市中西區健康路一段	22.98147	120.1978	Pizza,American	w60j
w60j_54826895.jpg	德式香腸披薩	54826895	11吋 整片30CM大	297	(△)臺南市中西區健康路一段	22.98147	120.1978	Pizza,American	w60j
w60j_54826896.jpg	起司起司披薩	54826896	11吋 整片30CM大	297	(△)臺南市中西區健康路一段	22.98147	120.1978	Pizza,American	w60j

清整完剩110萬筆餐點

- 價格清整
- 地址清整
- 類別清整

清整後

pic_id	food_name	food_id	introduce	price	address	latitude	longitude	label	shop_id
h2jy_16820023.jpg	雞肉飯	16820023	豬脂（油）原產地	36	臺南市安南區安和路	23.03077	120.2114	TW	h2jy
h2jy_16820024.jpg	魚鬆肉燥飯	16820024	肉品原產地：台灣	32	臺南市安南區安和路	23.03077	120.2114	TW	h2jy
h2jy_16820025.jpg	貢丸湯	16820025	肉品原產地：台灣	32	臺南市安南區安和路	23.03077	120.2114	TW	h2jy
h2jy_16820026.jpg	魚丸湯	16820026	豬脂（油）原產地	32	臺南市安南區安和路	23.03077	120.2114	TW	h2jy
w60j_54826892.jpg	夏威夷披薩	54826892	11吋 整片30CM大	297	臺南市中西區健康路	22.98147	120.1978	American	w60j
w60j_54826893.jpg	BBQ雞披薩	54826893	11吋 整片30CM大	297	臺南市中西區健康路	22.98147	120.1978	American	w60j
w60j_54826895.jpg	德式香腸披薩	54826895	11吋 整片30CM大	297	臺南市中西區健康路	22.98147	120.1978	American	w60j
w60j_54826896.jpg	起司起司披薩	54826896	11吋 整片30CM大	297	臺南市中西區健康路	22.98147	120.1978	American	w60j

早到晚到 彰化店



早到晚到 彰化店

4.7 ★★★★★ (55)

餐廳

總覽

評論

簡介



規劃路線



儲存



附近



傳送到手機



分享

✓ 內用 · ✓ 外帶 · ✓ 外送



500彰化縣彰化市中正路二段148號



營業中 · 關閉時間：01:30



預訂

04-722-0415

魏博十善生命館

附近的餐廳

飯店

觀光景點

酒吧

咖啡

外帶

超市



伊蝶汽車旅館

Riu美人館

章魚先生品味小館
墨西哥夾餅 · SS

總覽

評論

簡介

4.7

★★★★★

55 篇評論

撰寫評論



排序

全部

宵夜 4

炸 4

店員 3



賴賴生活圈

在地嚮導 · 74 則評論 · 1,585 張相片

★★★★★ 2 個月前

彰化新開的餐飲店，在火車站與員林客運之間。而且是一間主打晚餐跟宵夜的店。

彰化市的夜生活飲食，終於多一個美味的選擇了！！...

全文

易燥肉飯
料理 · S

地圖資料 ©2023

台灣

使用條款

隱私權 提供

網路爬蟲

Google Map

主講者：王柏崴

店家數變28783家
留言共363萬筆

店家ID	店名	評論	星星數
mq7f	早到晚到(彰化中正店)	東西超好吃 份量又大 店員人也超親切人很好  泡麵附ㄉ蛋還是半熟的..	5 顆星
mq7f	早到晚到(彰化中正店)	彰化新開的餐飲店，在火車站與員林客運之間。而且是一間主打晚餐跟宵夜的店。 彰化市的夜生活飲食，終於多一個美味的選擇了！！	5 顆星

- 用selenium套入第一表單的店家名稱進行搜尋
- 排除非唯一店家名稱

資料清整

Google Map

原始資料

店家ID	店名	評論	星星數
mq7f	早到晚到(彰化中正店)	東西超好吃 份量又大 店員人也超親切人很好 	5 頤星
mq7f	早到晚到(彰化中正店)	彰化新開的餐飲店，在火車站與員林客運之間。而且是一間主打晚餐跟宵夜的店。 彰化市的夜生活飲食，終於多一個美味的選擇了！！	5 頤星

結巴+刪除停用字

店家ID	店名	評論	星星數
mq7f	早到晚到(彰化中正店)	東西超好吃 份量店員人超親切 泡麵	5 頤星
mq7f	早到晚到(彰化中正店)	彰化 新開 餐飲店 火車站 與員林客運	5 頤星
mq7f	早到晚到(彰化中正店)	東西 好吃 好吃 黑輪片 炸黑 可怕	2 頤星
mq7f	早到晚到(彰化中正店)	其實 不用 賣 多品 項光 炒泡麵 稱霸	4 頤星
mq7f	早到晚到(彰化中正店)	uber 上點 一下子 餐點 之後對 豬排 咖哩	5 頤星
mq7f	早到晚到(彰化中正店)	服務 態度 餐點 不錯 可能 剛開幕 人手少	5 頤星
mq7f	早到晚到(彰化中正店)	店點 餐 告知 掃 qr code 點餐 掃 發現 要	1 頤星
mq7f	早到晚到(彰化中正店)	好吃 胖闆娘 親切 漂亮	5 頤星
mq7f	早到晚到(彰化中正店)	1029 晚上 646 730 餐點 全到 齊 內用 一	3 頤星
mq7f	早到晚到(彰化中正店)	餐點 好吃 服務 態度 良好	5 頤星
mq7f	早到晚到(彰化中正店)	不錯 快吃	5 頤星

- 特色分類
 - ✓ 服務
 - ✓ 環境
 - ✓ 美味

特色歸類

StoreID	StoreName	Label
mq7f	早到晚到(彰化中正店)	服務
mqeo	五美茶果茶(嘉義民族店)	服務
mqm2	紅勁冰品茶飲(彰化花壇店)	環境
mqr7	鮮茶道(彰化田中店)	美味

資料合併

1 2 3

pic_id	food_name	food_id	introduce	address	latitude	longitude	label	shop_id	sort	price	span
mq7f_96743787.jpg	(F)手感豬肉漢堡	96743787	肉品原產地：台	彰化縣彰化市中正路二段	24.07873	120.5374	Brunch	mq7f	服務	65	2
mq7f_96743788.jpg	(F)法式牛奶起司雞	96743788		彰化縣彰化市中正路二段	24.07873	120.5374	Brunch	mq7f	服務	188	4
mq7f_96743789.jpg	花生厚片	96743789	香濃滑順的花生	彰化縣彰化市中正路二段	24.07873	120.5374	Brunch	mq7f	服務	65	2
mq7f_96743791.jpg	(F)濃厚乳酪花生法	96743791	香濃滑順的花生	彰化縣彰化市中正路二段	24.07873	120.5374	Brunch	mq7f	服務	80	2
mq7f_96743792.jpg	薯泥起司蛋餅	96743792	將綿密薯泥、濃	彰化縣彰化市中正路二段	24.07873	120.5374	Brunch	mq7f	服務	75	2
mq7f_96743793.jpg	(F)雙層牛肉起士堡	96743793		彰化縣彰化市中正路二段	24.07873	120.5374	Brunch	mq7f	服務	154	4
mq7f_96743796.jpg	(F)愜意吃爆乳芋泥	96743796		彰化縣彰化市中正路二段	24.07873	120.5374	Brunch	mq7f	服務	100	2
mq7f_96743797.jpg	(F)肉蛋起司瀑布三	96743797		彰化縣彰化市中正路二段	24.07873	120.5374	Brunch	mq7f	服務	110	3
mq7f_96743867.jpg	(F)卡滋薯餅蛋堡	96743867	New	彰化縣彰化市中正路二段	24.07873	120.5374	Brunch	mq7f	服務	74	2
mq7f_96743868.jpg	(F)特製雙層手感豬	96743868	肉品原產地：台	彰化縣彰化市中正路二段	24.07873	120.5374	Brunch	mq7f	服務	87	2
mq7f_96743869.jpg	(F)巧克力卡啦漢堡	96743869		彰化縣彰化市中正路二段	24.07873	120.5374	Brunch	mq7f	服務	94	2
mq7f_96743870.jpg	(F)巧克力卡啦漢堡	96743870		彰化縣彰化市中正路二段	24.07873	120.5374	Brunch	mq7f	服務	94	2
mq7f_96743871.jpg	(F)卡啦不辣漢堡	96743871		彰化縣彰化市中正路二段	24.07873	120.5374	Brunch	mq7f	服務	94	2
mq7f_96743873.jpg	(F)朱熹雙層燒肉堡	96743873		彰化縣彰化市中正路二段	24.07873	120.5374	Brunch	mq7f	服務	107	3
mq7f_96743875.jpg	(F)阿嬤養的漢堡豬	96743875	肉品原產地：台	彰化縣彰化市中正路二段	24.07873	120.5374	Brunch	mq7f	服務	127	3
mq7f_96743876.jpg	(F)霸王雞排堡	96743876		彰化縣彰化市中正路二段	24.07873	120.5374	Brunch	mq7f	服務	127	3
mq7f_96743857.jpg	(F)原味鍋燒麵	96743857	麵類擇一	彰化縣彰化市中正路二段	24.07873	120.5374	Brunch	mq7f	服務	127	3
mq7f_96743858.jpg	(F)韓式泡菜鍋燒麵	96743858	麵類擇一	彰化縣彰化市中正路二段	24.07873	120.5374	Brunch	mq7f	服務	134	3
mq7f_96743859.jpg	(F)日式咖哩鍋燒烏	96743859		彰化縣彰化市中正路二段	24.07873	120.5374	Brunch	mq7f	服務	154	4
mq7f_96743799.jpg	(F)就是原味雙蛋蛋	96743799		彰化縣彰化市中正路二段	24.07873	120.5374	Brunch	mq7f	服務	65	2
mq7f_96743800.jpg	九層塔蛋餅	96743800	香氣濃郁的九層	彰化縣彰化市中正路二段	24.07873	120.5374	Brunch	mq7f	服務	65	2

4. 比對餐點圖片資料夾，互不存在刪除
5. 圖片處理(像素轉換/黑色矩陣)



System

系統工程

Manage

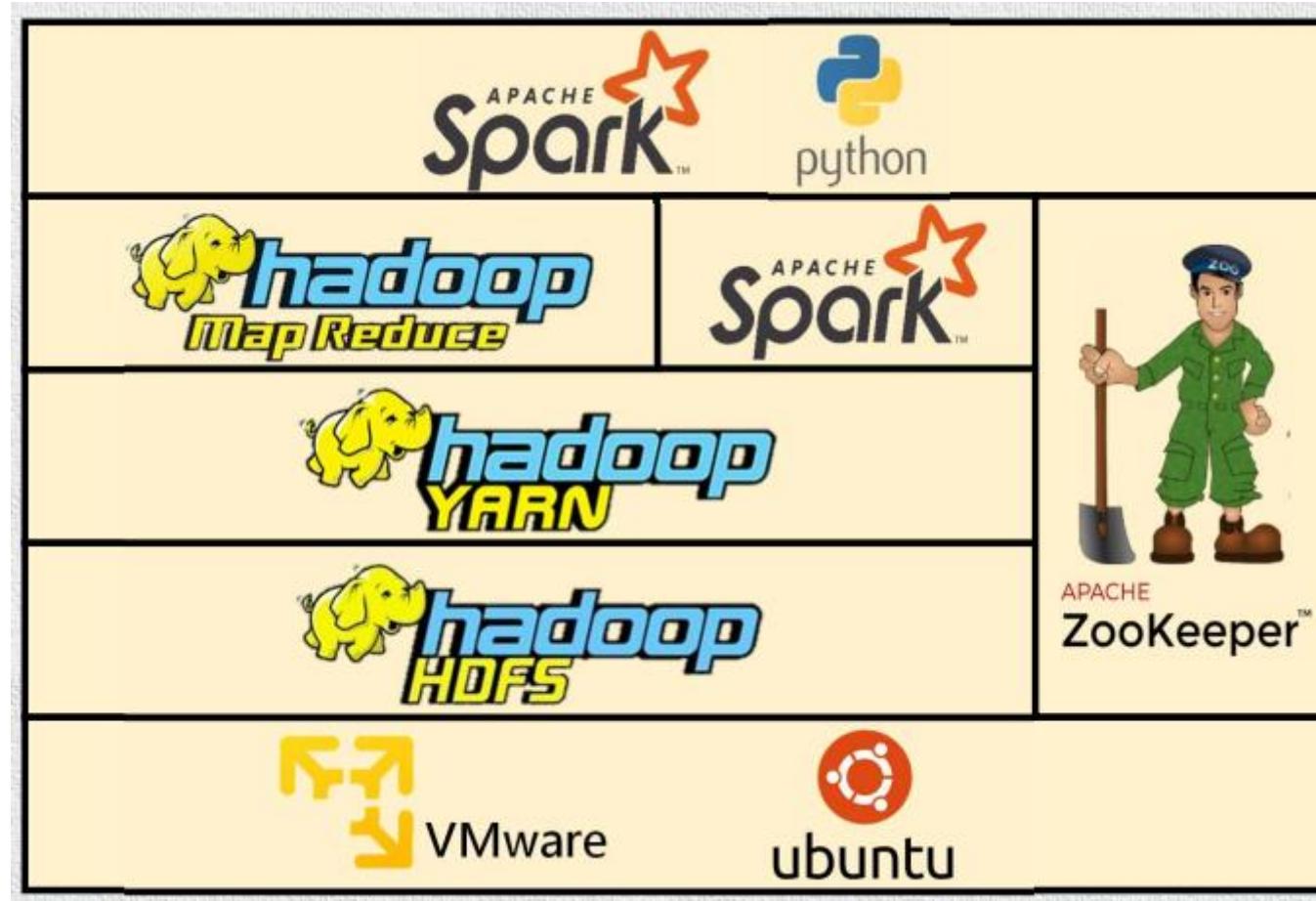
資料管理



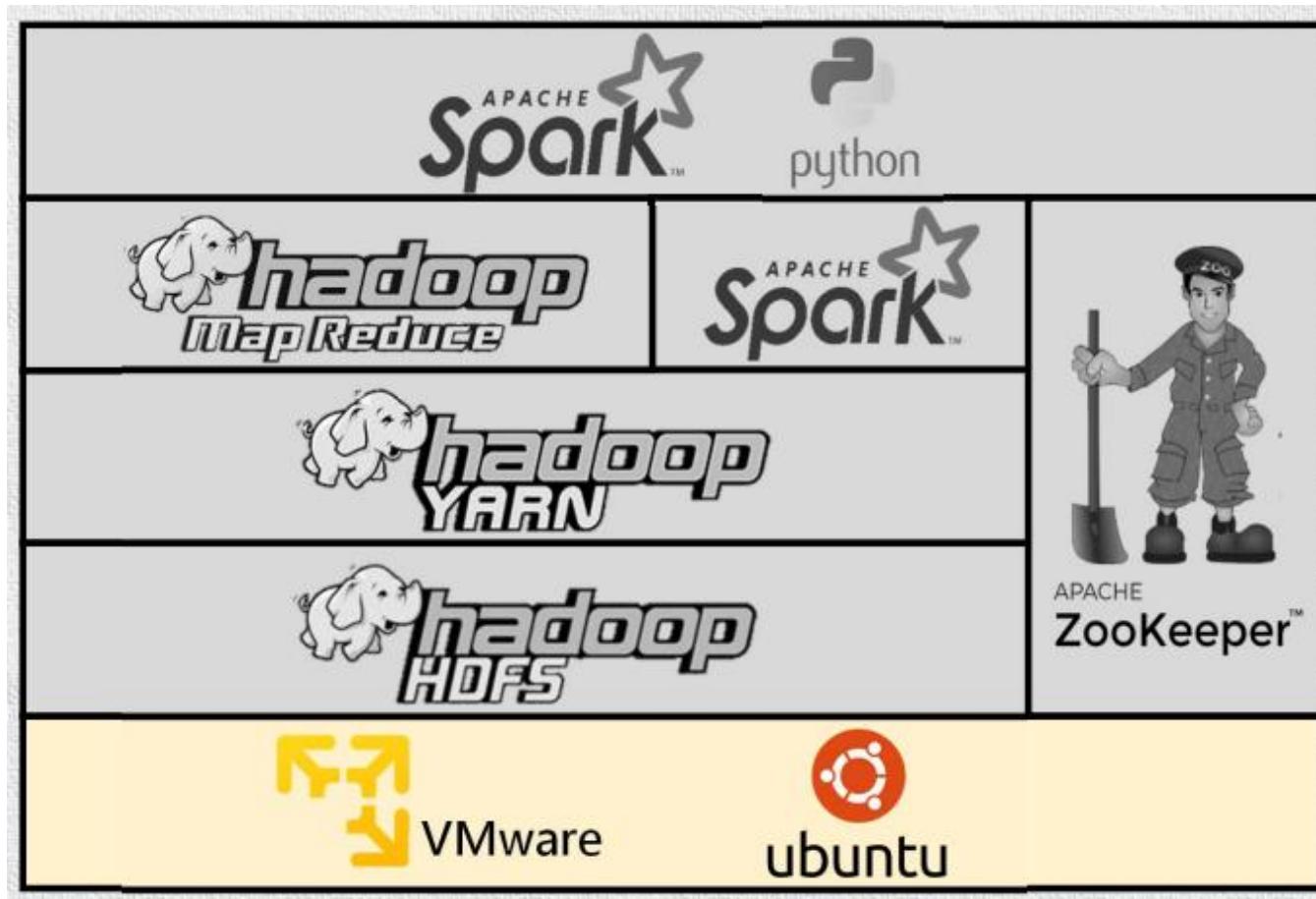
黃湘玲

Speaker

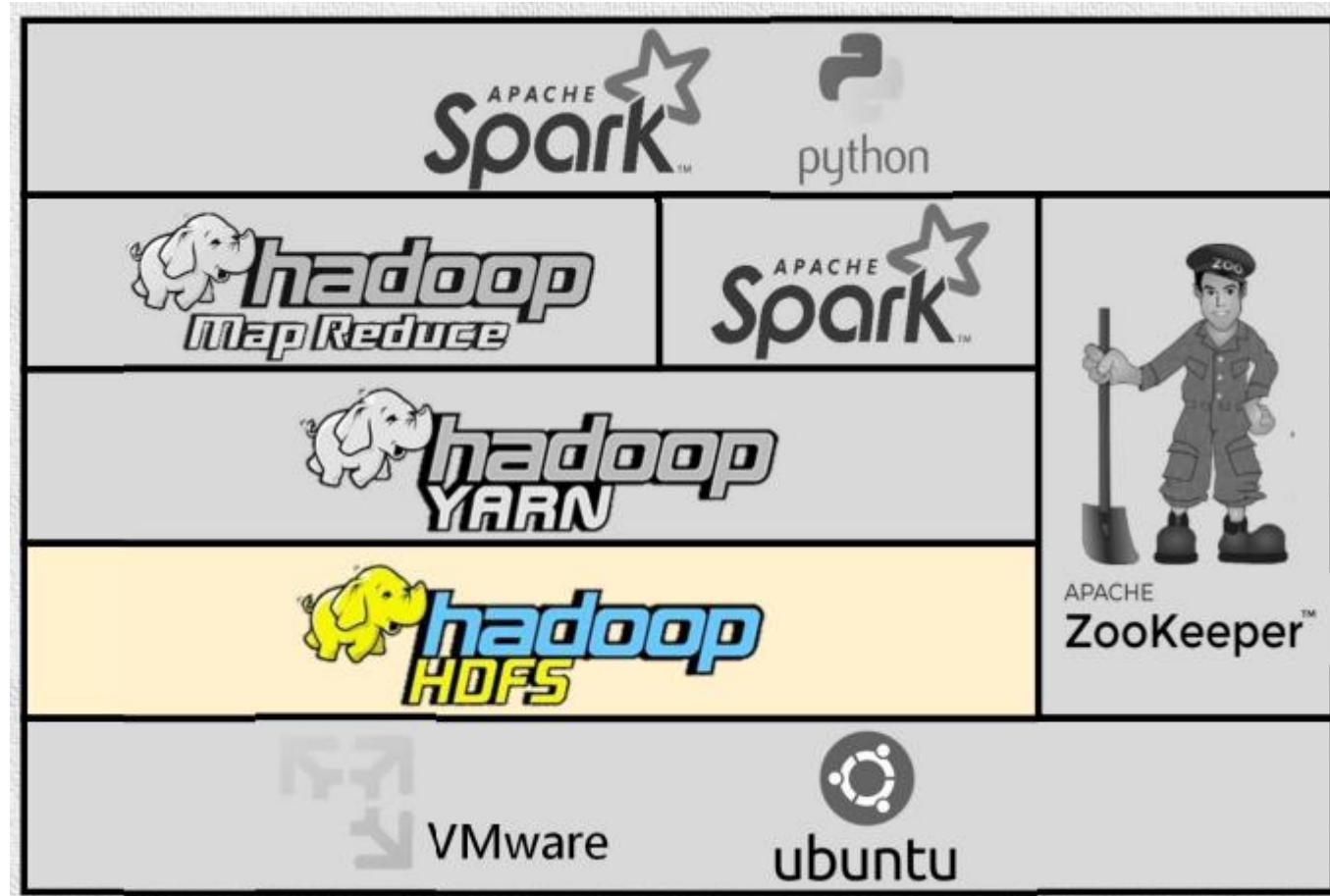
環境建置-架構



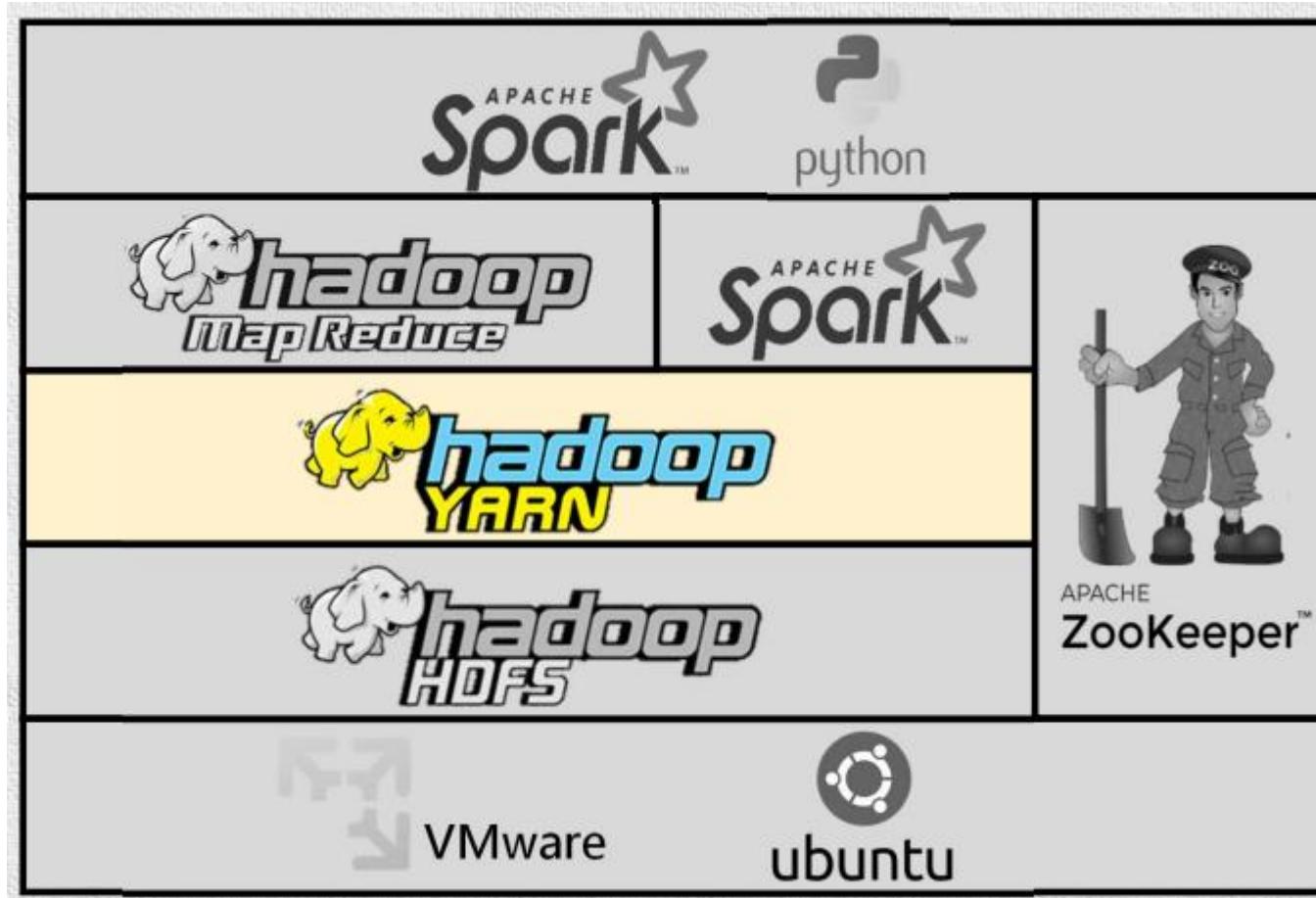
環境建置-作業系統



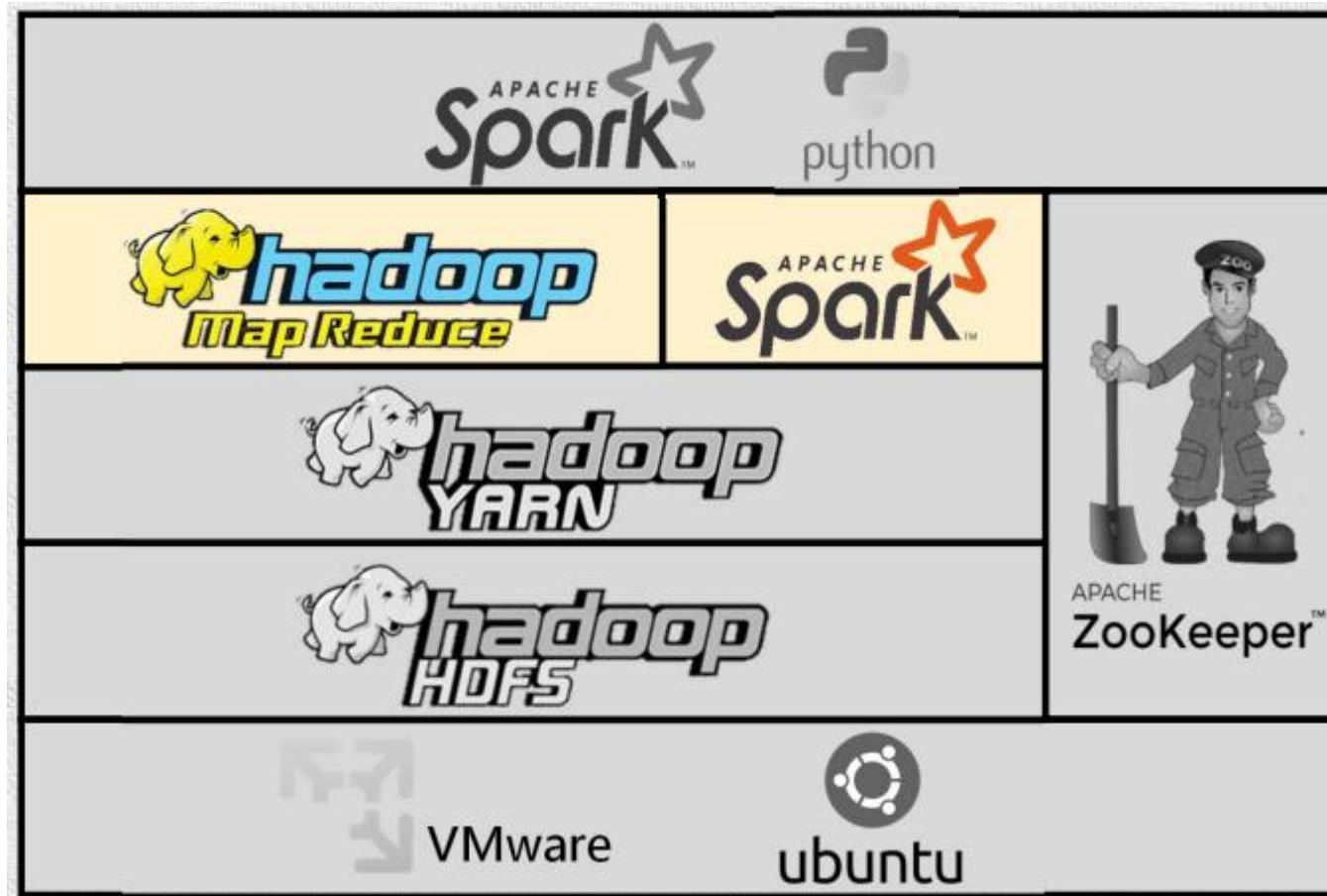
環境建置-HDFS



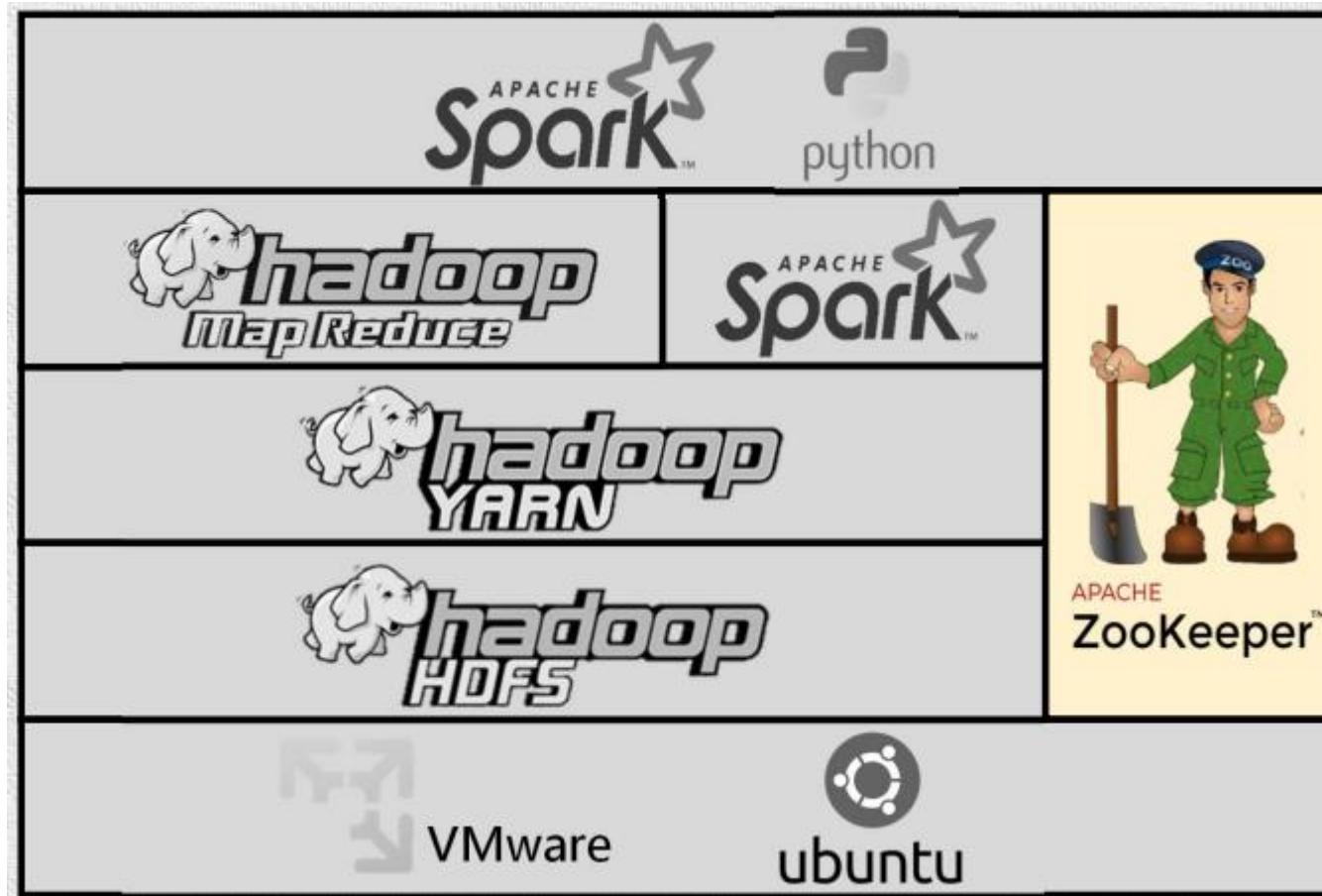
環境建置-YARN



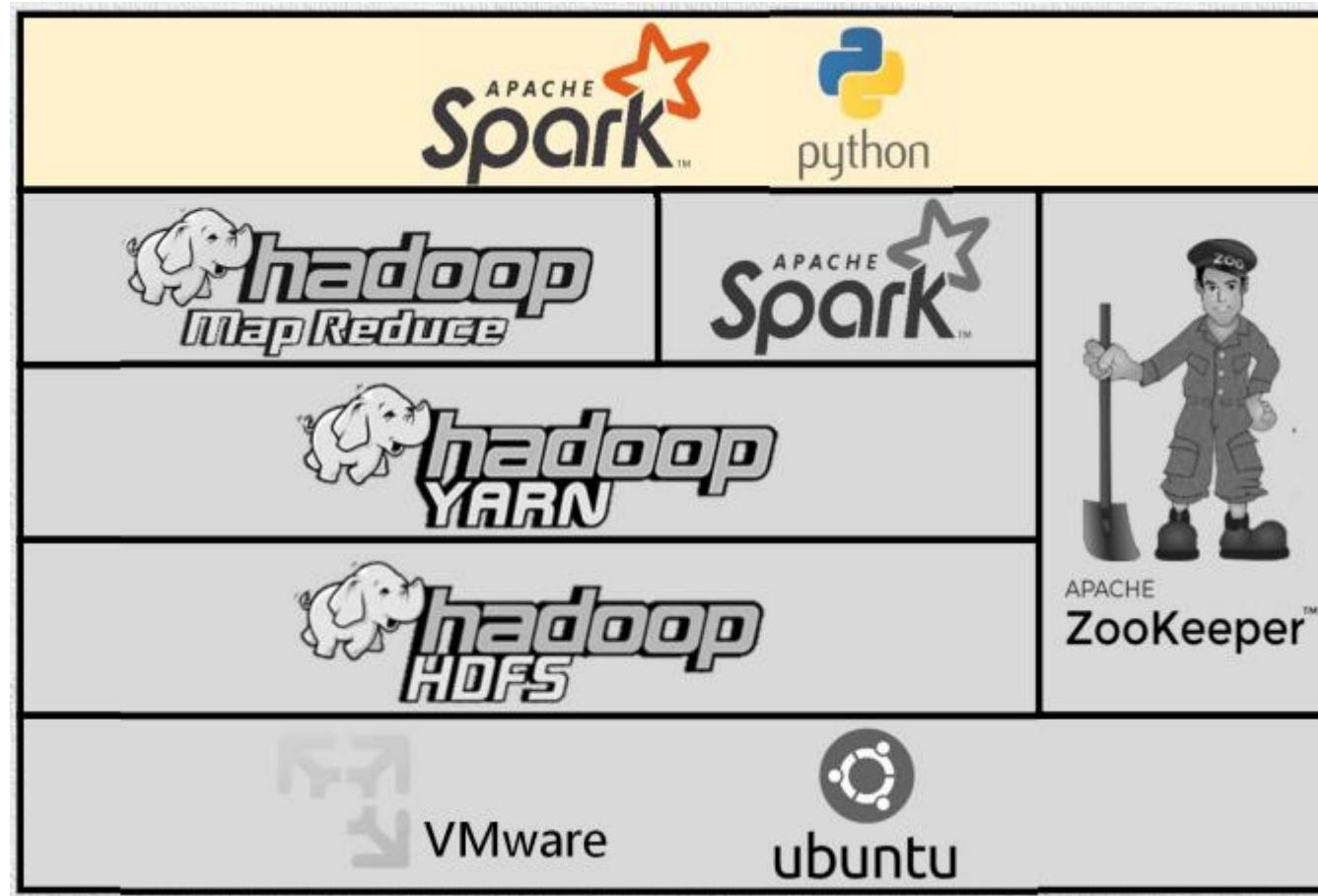
環境建置 -MapReduce/RDD



環境建置-ZooKeeper

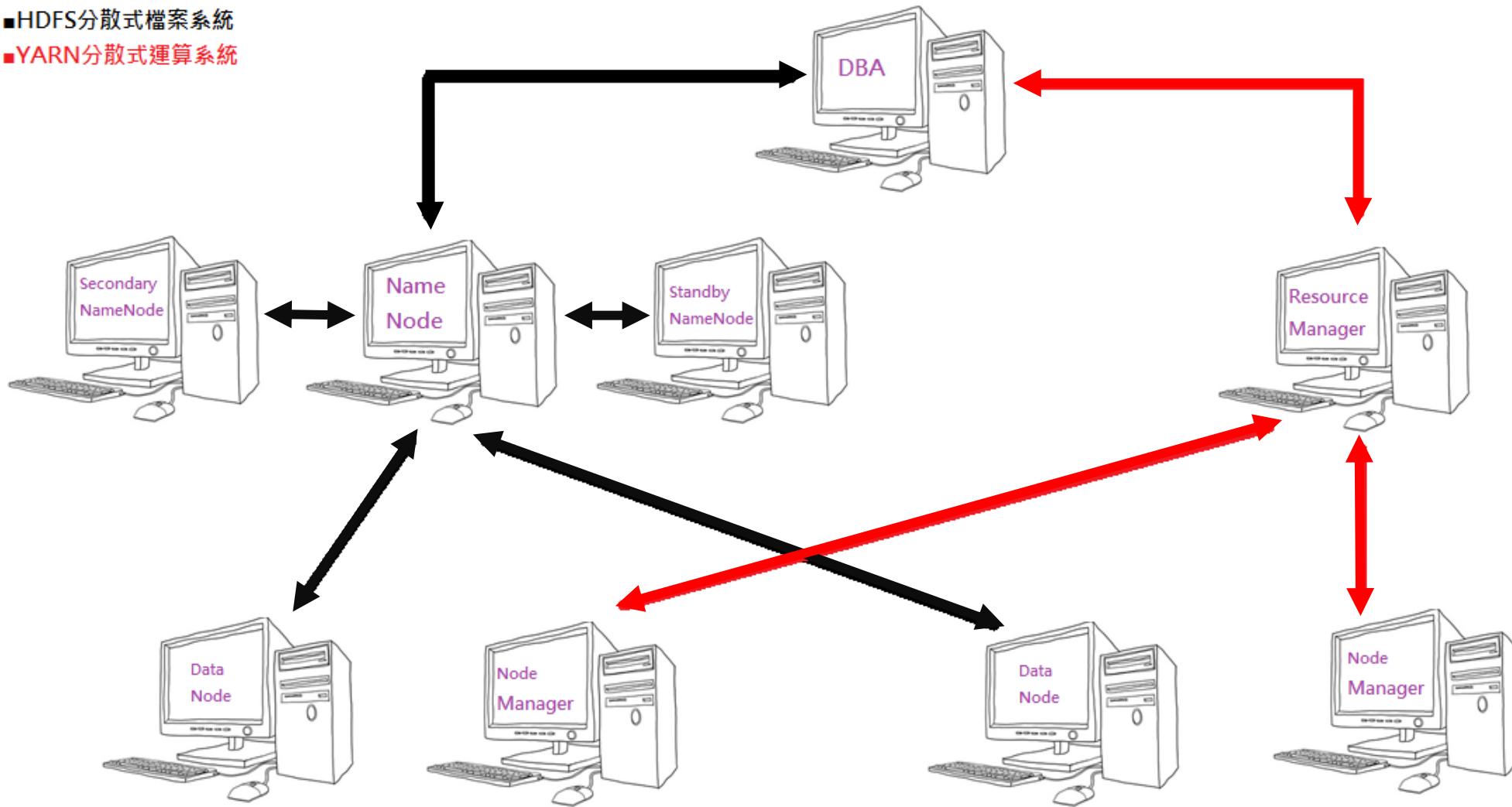


環境建置 -Spark / python



Hadoop運作方式

- HDFS分散式檔案系統
- YARN分散式運算系統



資源配置

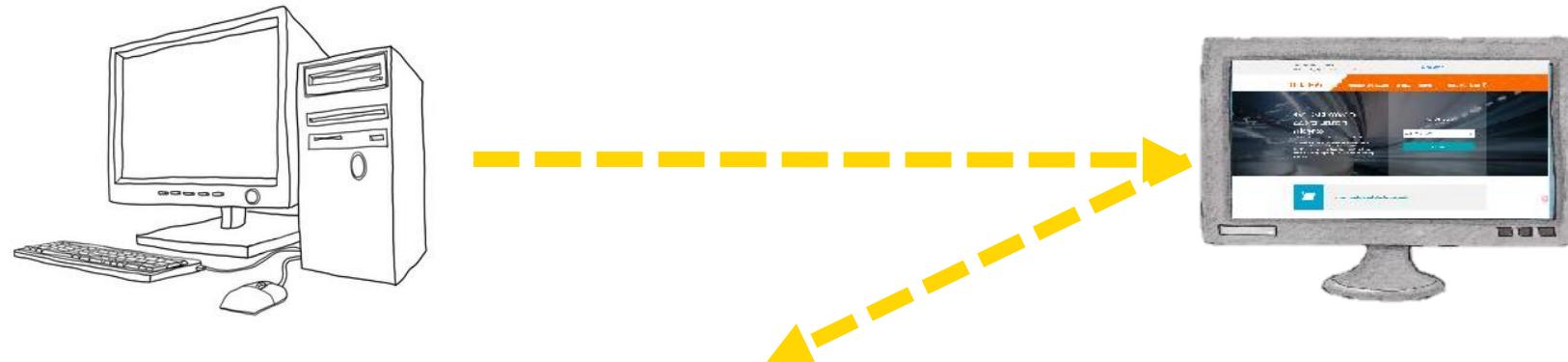
名稱	功能	說明	記憶體	CPU
DBA	最高管理權限	用於管理HDFS、YARN、SQL和Worker	16G	3核
Name Node	HDFS	一種分散式檔案系統，提供對應用程式資料的高吞吐量存取	16G	3核
Resource Manager	YARN	作業排程與叢集資源管理的框架以及Map Reduce用於平行運算處理	16G	3核
Hive Server 2	SQL	用於存放Hadoop中的資料集	16G	3核
worker	worker	處理效能	320G	60核

共8台實體電腦

每台實體電腦配置3台虛擬電腦

每台虛擬電腦配置16G記憶體和3顆CPU

如何用tableau串接Hive Server 2



hive_jdbc_2.6.5.1007.zip



Tableau串接



Efficiency

效能檢測

Analysis

資料分析



郭安迪
Speaker

效能檢測

Setting Parameters

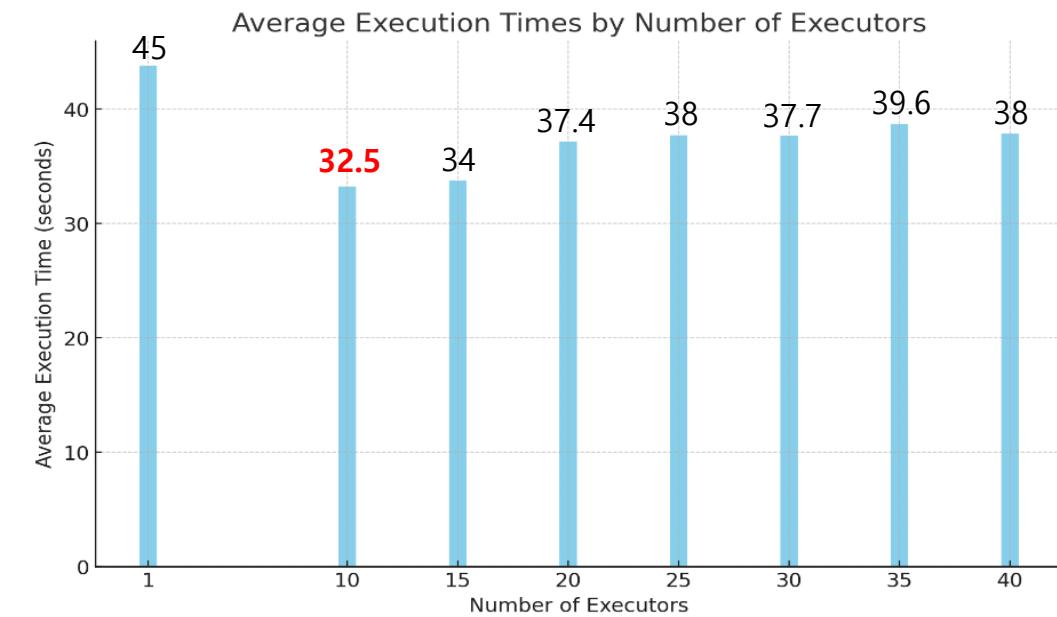
參數	參數值	說明
Executor	40	With 20 virtual machine, we can use up to 40 executors.
Core	1, 2, 3	We testing with 1-3 core to decide which one runs the fastest.
Memory (G)	5G	Maximum capacity for spark executor memory can up to 5G.

Executor
Core: 3
Memory: 5G



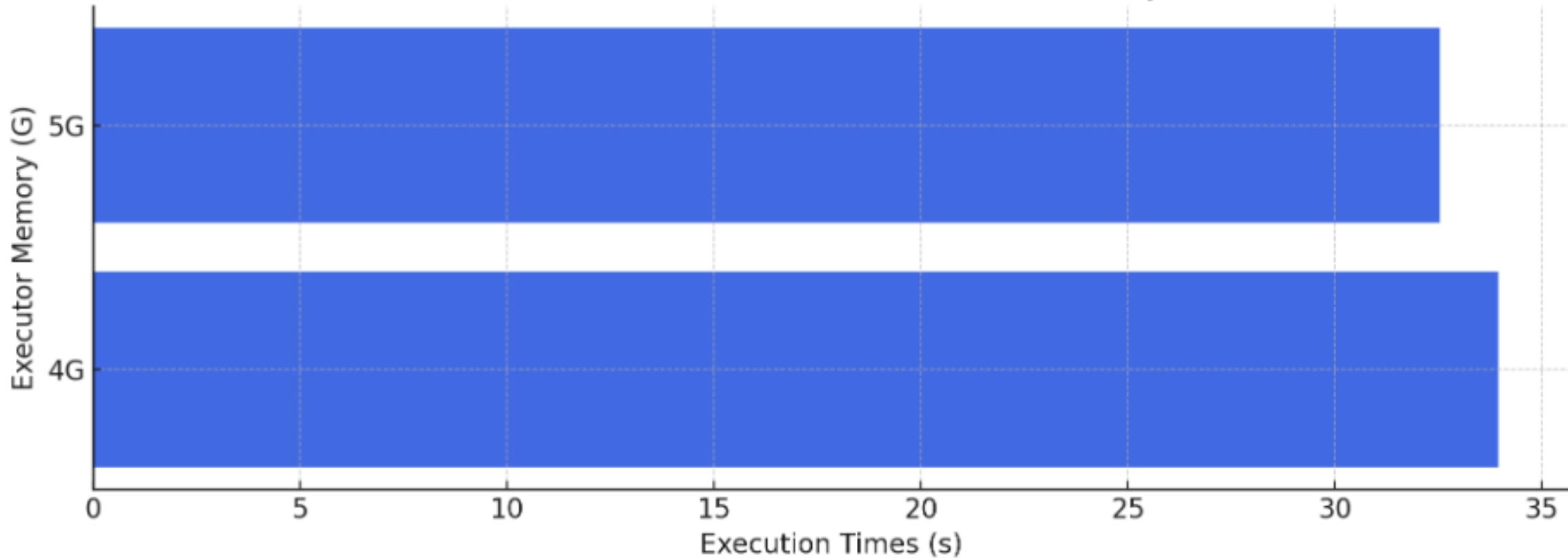
Testing Parameters

Executors	Cores	Memory (G)	Execution Times (s)
1	3	5	45.0
10	3	5	32.5
15	3	5	34.0
20	3	5	37.4
25	3	4	38.0
30	3	4	37.7
35	3	4	39.6
40	3	4	37.9



效能檢測

Executors Memory comparisons



效能檢測

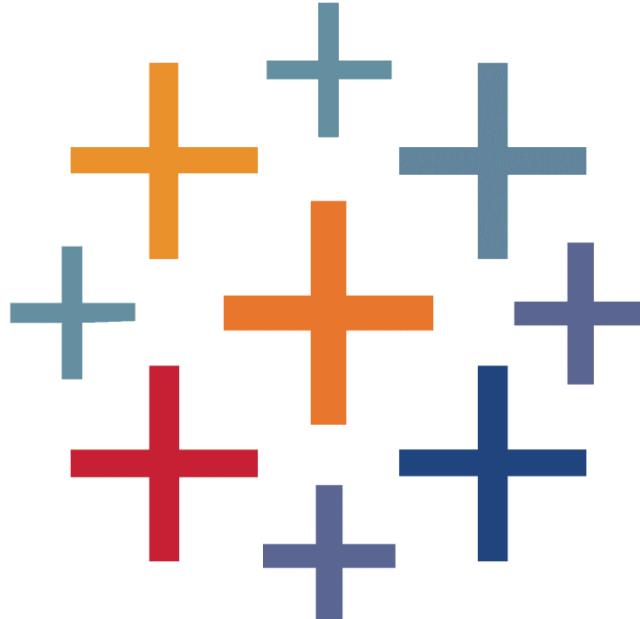
Result

環境	單機	叢集
訓練時間	45.03 秒	32.54 秒

{ 快13秒 }

使用工具

+ a b l e a u

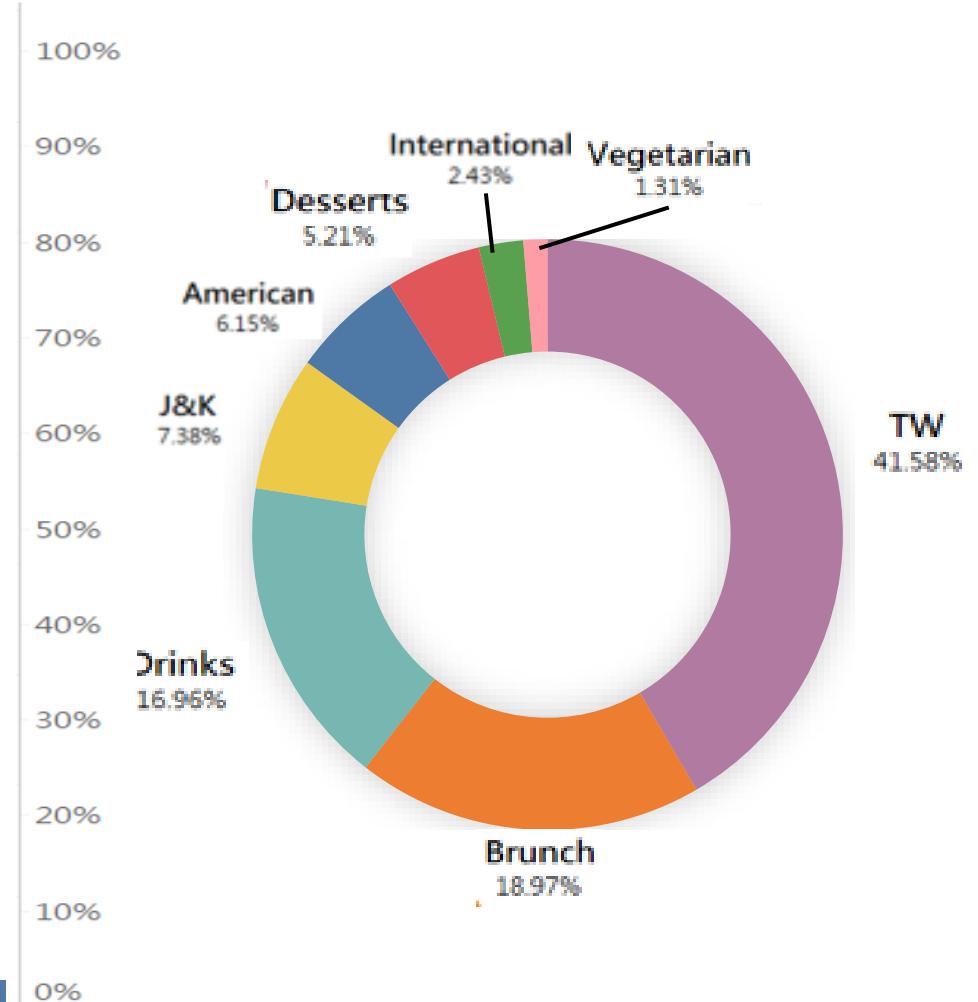
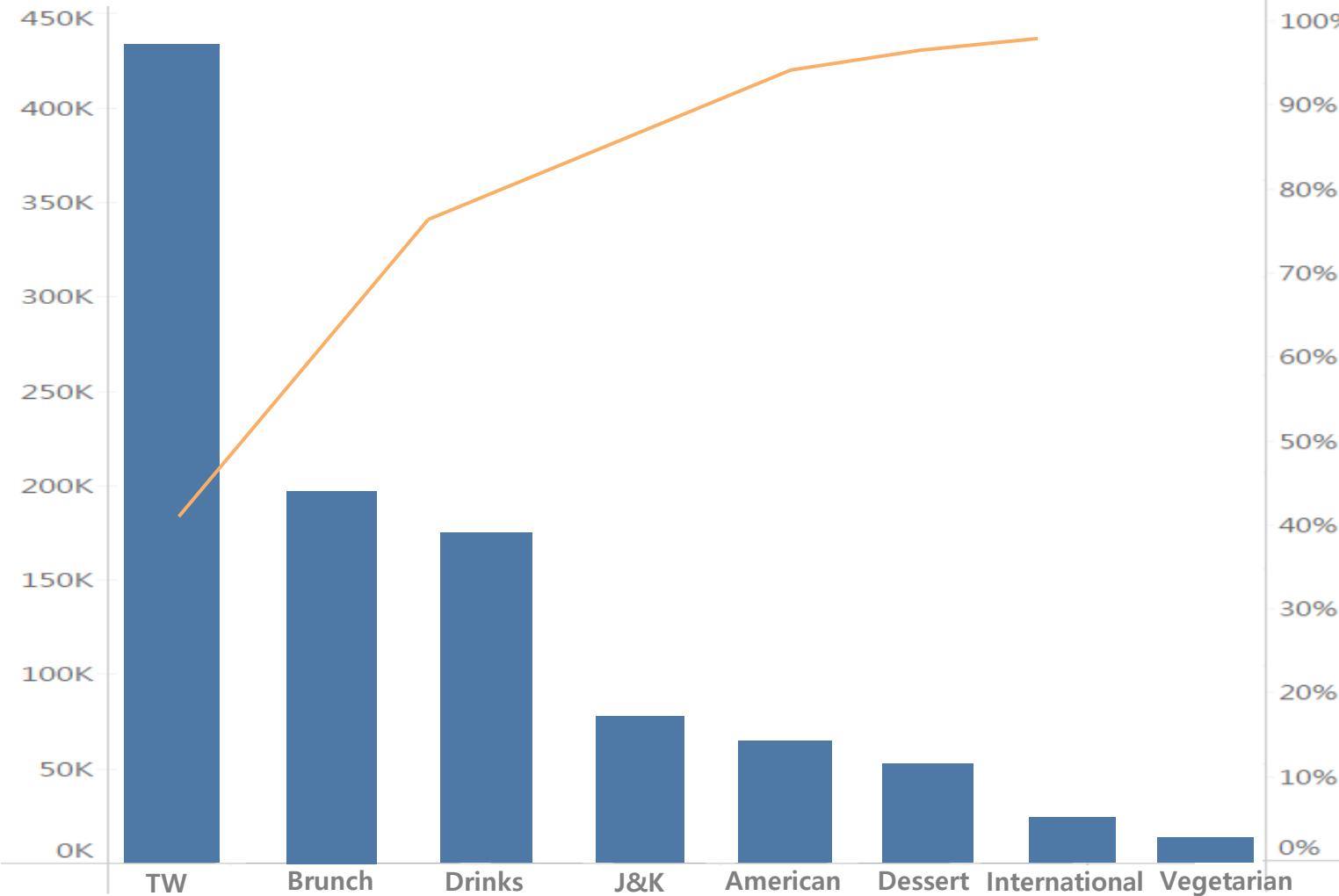


資料視覺化

各類別餐點分佈

主講者：郭安迪

各類餐點數量

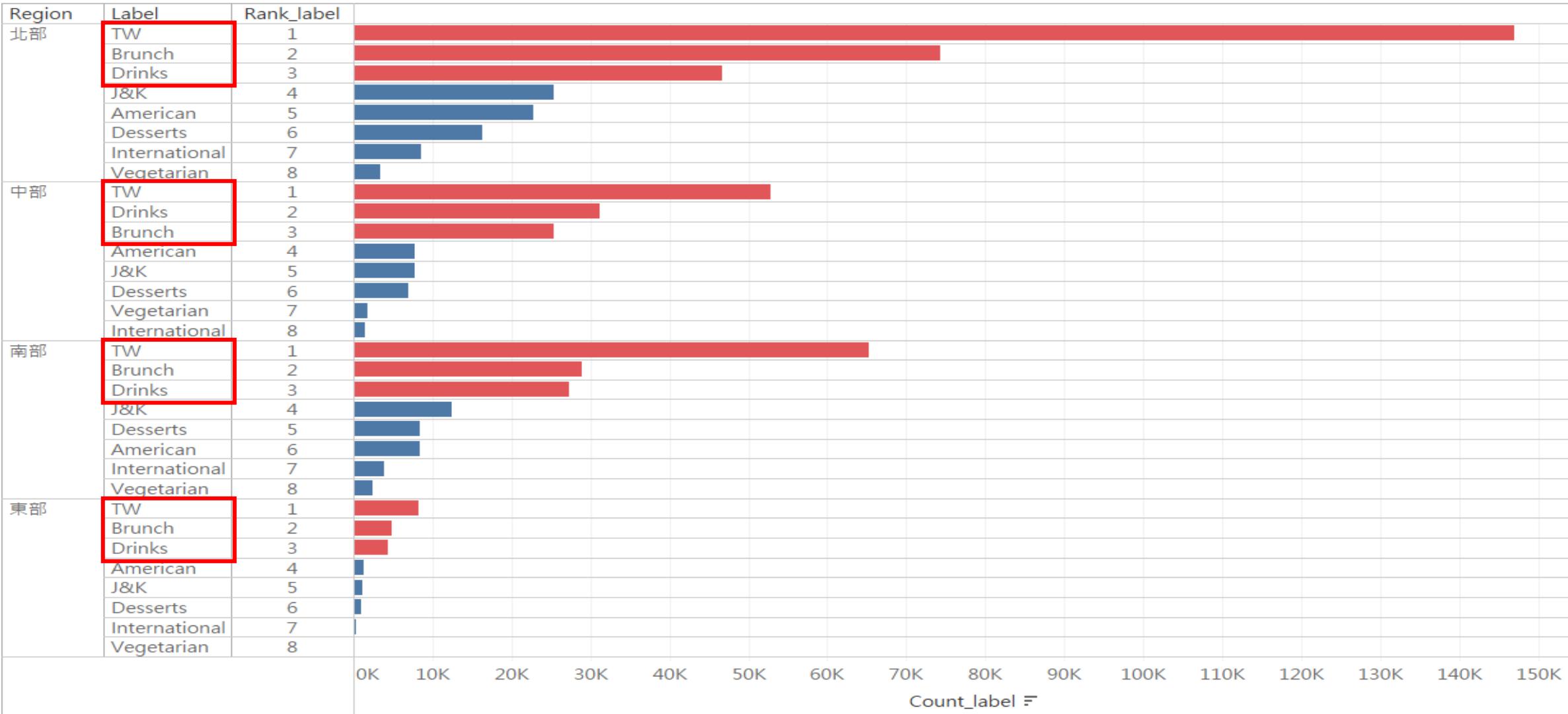


資料視覺化

各縣市Top3店家數量

主講者：郭安迪

各區Top3類別餐點

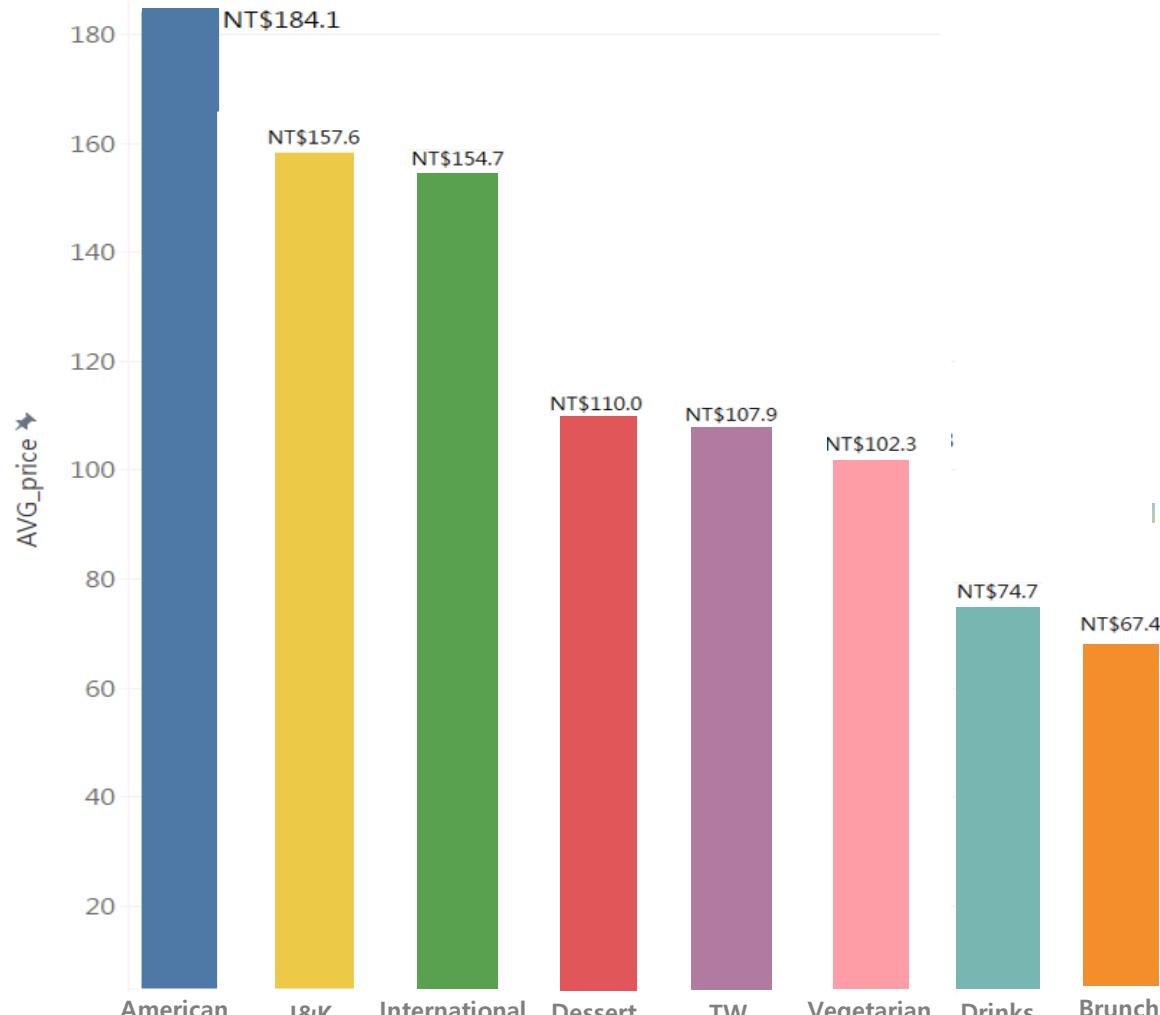


每個 Rank_label 的按 Region 與 Label 細分的 Count_label。色彩顯示有關 Rank_label_Color 的詳細資訊。檢視按 Region 進行篩選，這會排除 Null 與 福建省。

資料視覺化

各類餐點平均價格

各類餐點價格平均



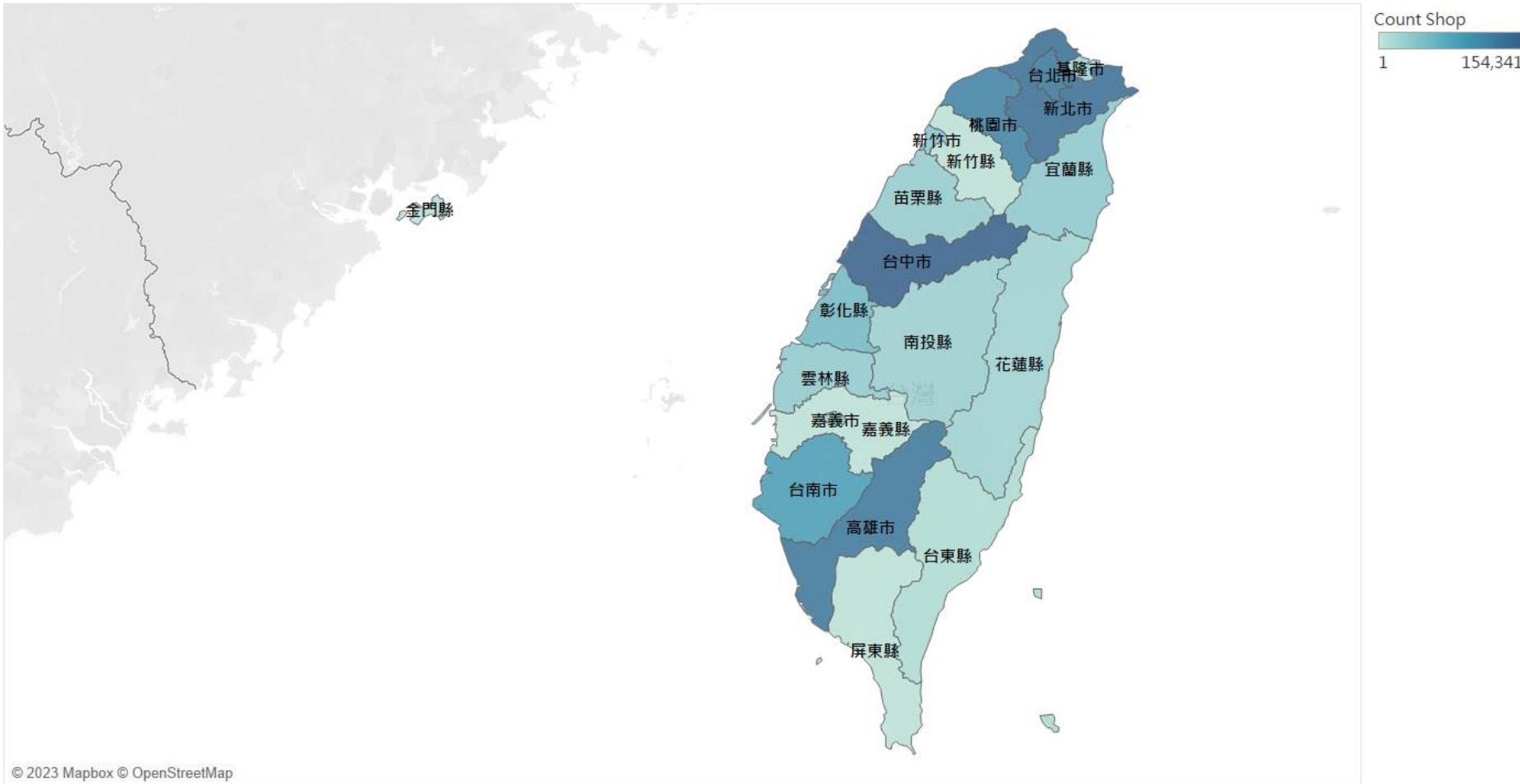
每個 Label 的 AVG_price。色彩顯示有關 Label 的詳細資訊。標記按 AVG_price
進行標示。

資料視覺化

各縣市店家總數

主講者：郭安迪

各縣市店家數量

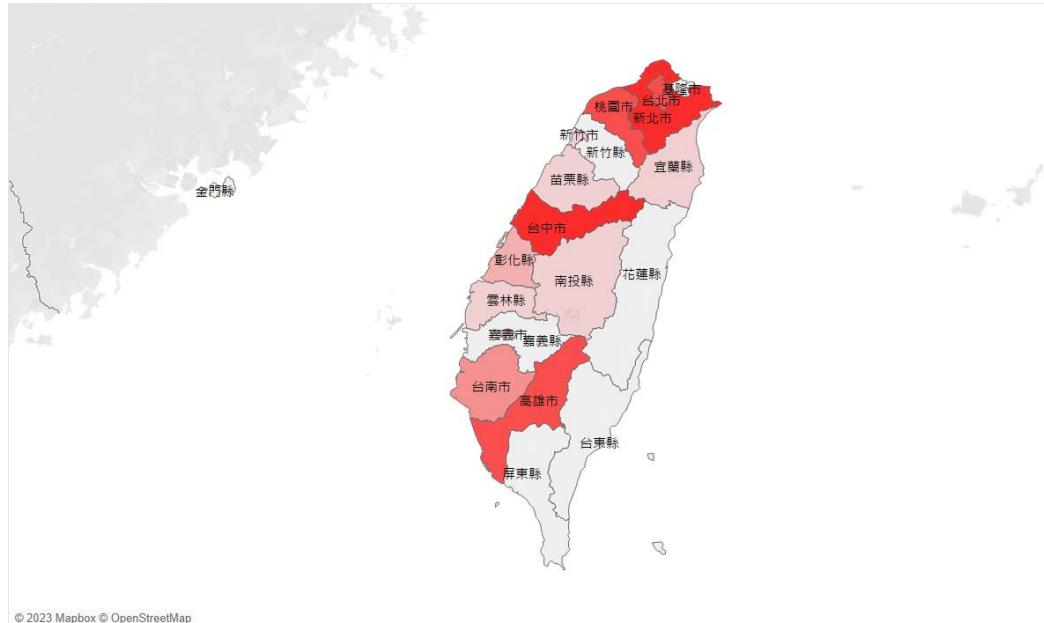


依據 經度(產生) 和 緯度(產生) 的地圖。 色彩顯示 Count Shop。 標記按 City (Taiwan Cities Coordinates.Csv) 進行標示。 為 City (Taiwan Cities Coordinates.Csv) 顯示詳細資訊。

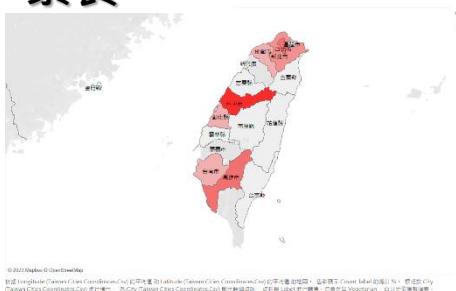
資料視覺化

各縣市店家密集度

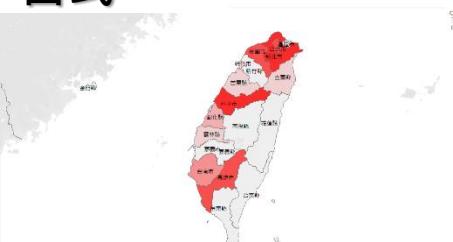
總計



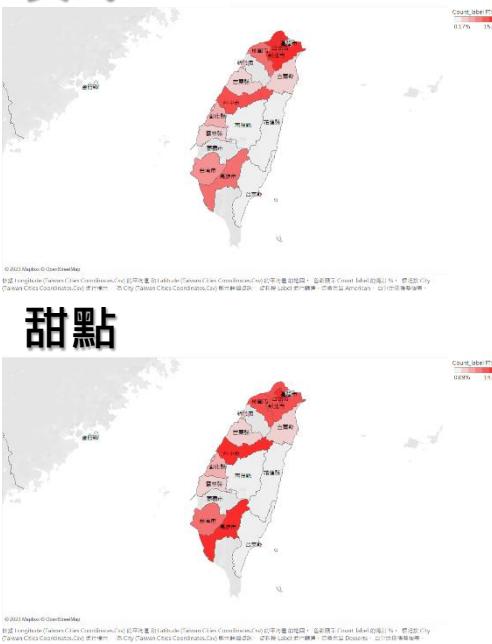
素食



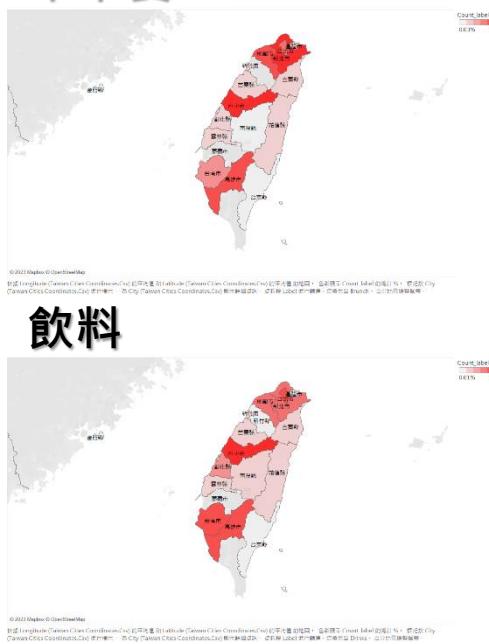
台式



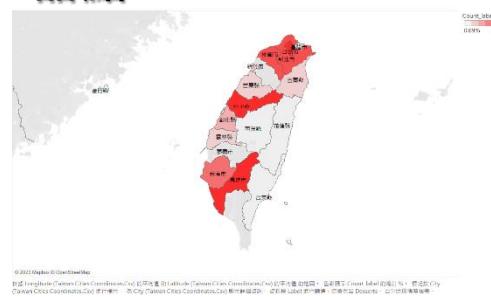
美式



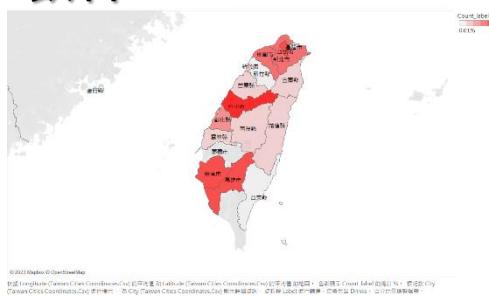
早午餐



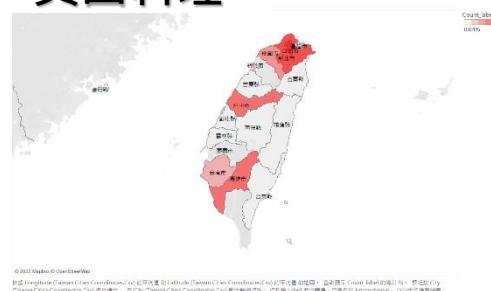
甜點



飲料



異國料理



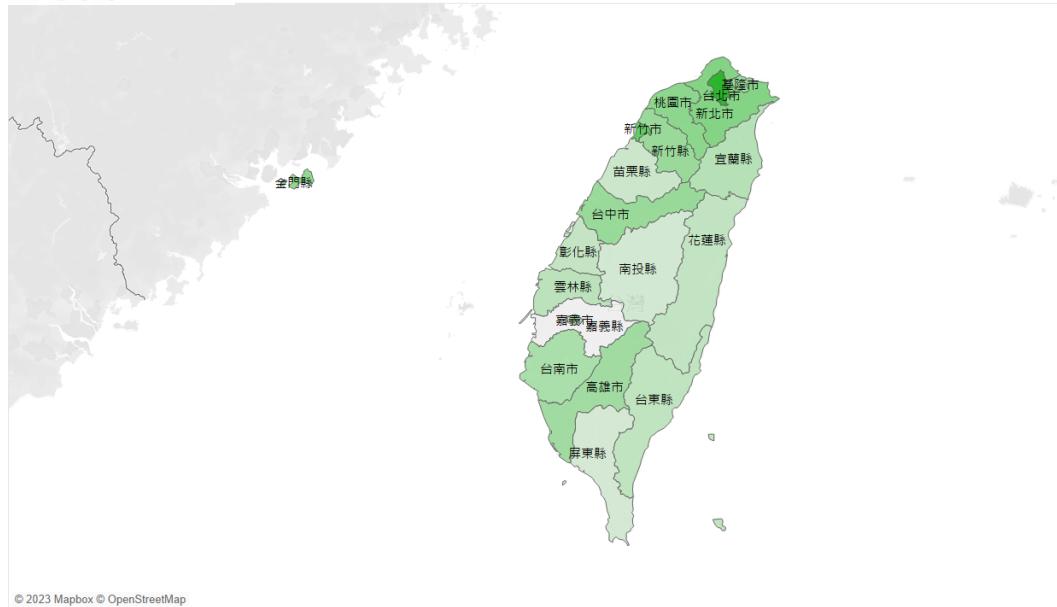
日韓



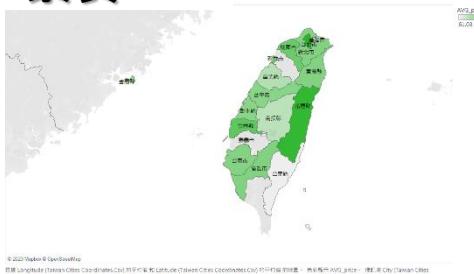
資料視覺化

各縣市平均價位

總計



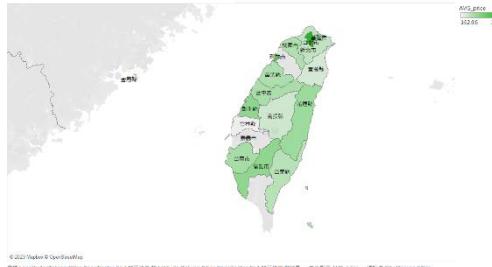
素食



台式



美式



早午餐



甜點



飲料



異國料理



日韓



資料視覺化

各區餐點平均價位

主講者：郭安迪

各區域的各類餐點價格分界

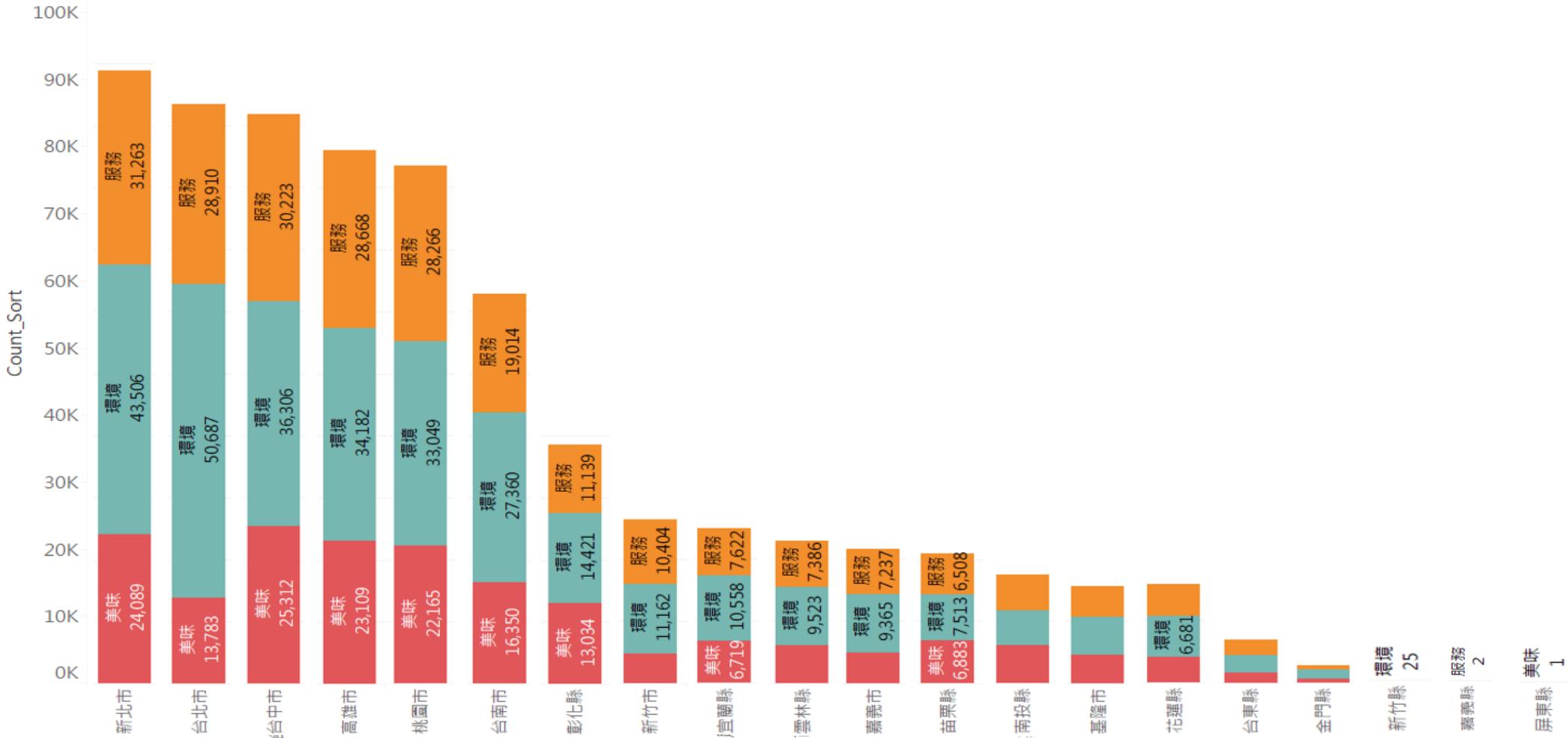


按 Region 以及 Label 細分的 AVG_price。 圖形顯示有關 Boolean_price 的詳細資訊。 標記按 AVG_price 進行標示。

資料視覺化

各縣市餐廳品質分析

各縣市餐廳品質



每個 City (Taiwan Cities Coordinates.Csv) 的 Count_Sort。色彩顯示有關 Sort 的詳細資訊。標記按 Sort 與 Count_Sort 進行標示。在 Sort 與 City (Taiwan Cities Coordinates.Csv) 上篩選檢視。Sort 篩選條件排除 Null。City (Taiwan Cities Coordinates.Csv) 篩選條件排除 Null。

Feature

特徵工程

Model

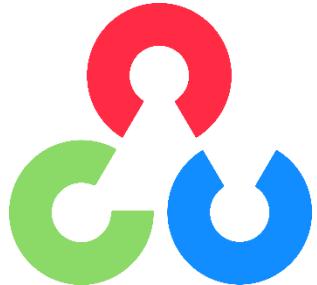
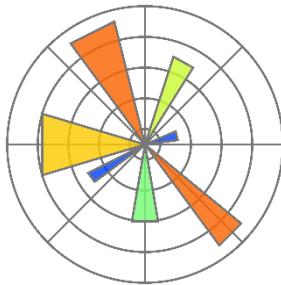
模型建置



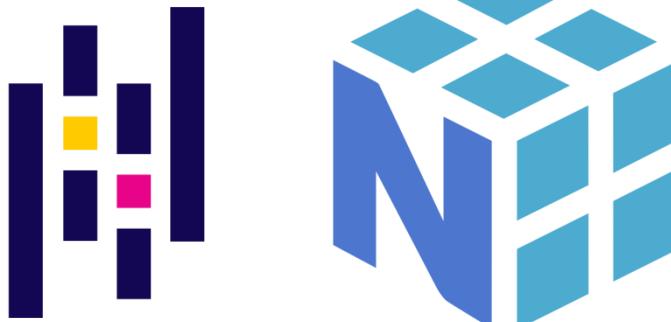
魯又齊
Speaker

使用工具

圖像處理與視覺化



資料處理



機器學習



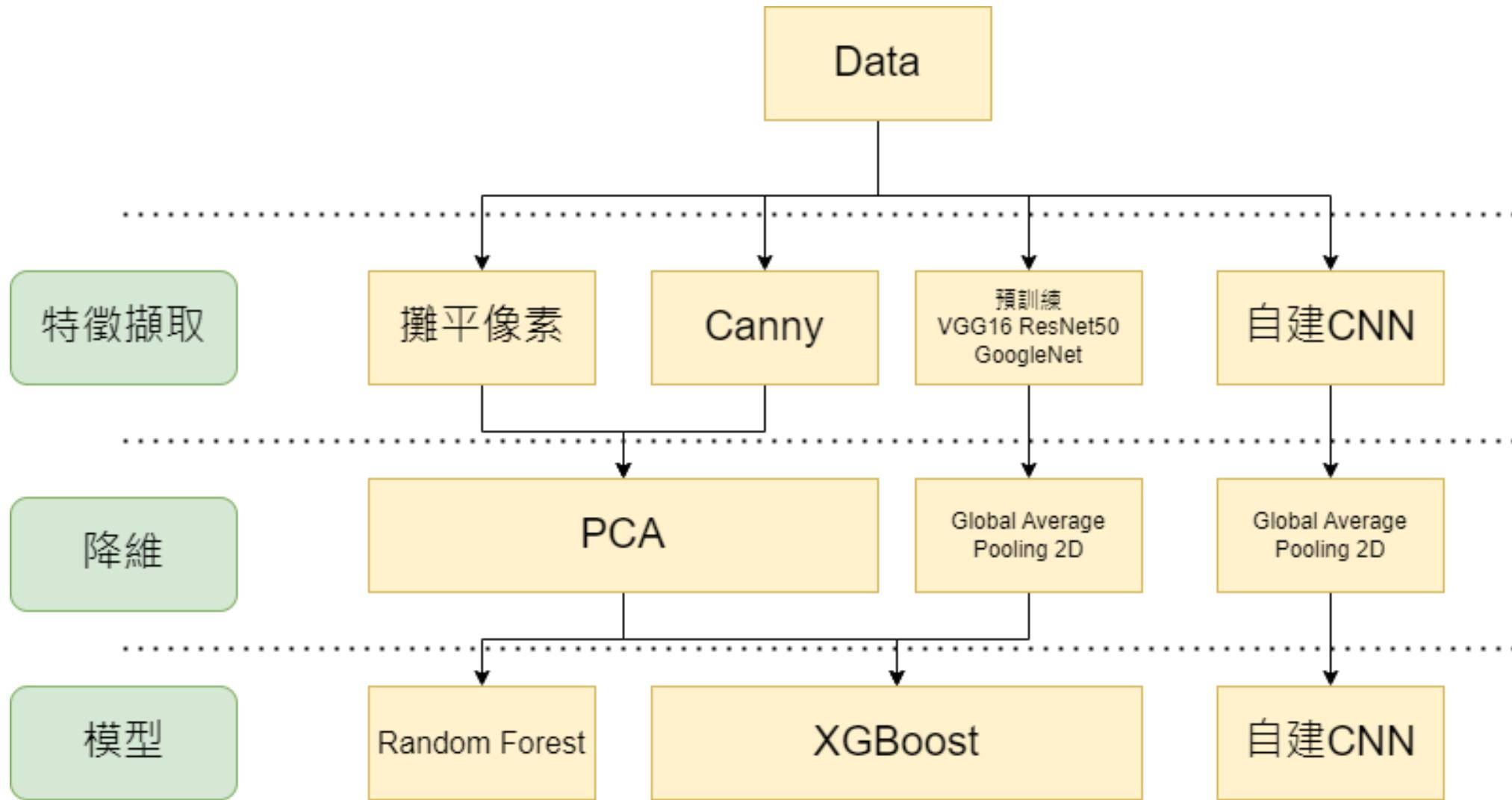
深度學習



環境

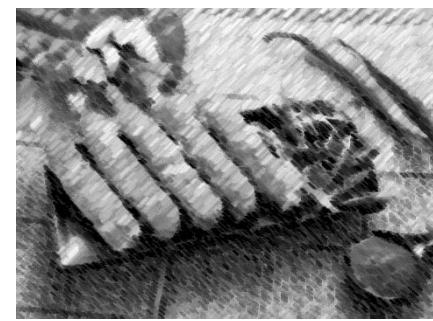


實驗地圖



圖片特徵擷取方式(1)

- 圖片轉換為128x128像素的灰階影像後攤平。



104, 83, 74, 122, 119, 80, 250, 78,
216, 244, 166, 249, 59, 160, 182, 1:
73, 132, 212, 160, 95, 36, 11, 92, :
243, 91, 183, 254, 19, 168, 158, 232
56, 12, 159, 210, 46, 182, 216, 208
49, 204, 120, 34, 1, 173, 25, 18,
184, 77, 230, 223, 162, 193, 215, 22
140, 119, 206, 22, 3, 107, 143, 16:
122, 199, 192, 71, 33, 209, 43, 242
181, 135, 106, 96, 205, 253, 195, 58,



, 57, 123, 42, 235, 181, 39, 114, 168, 166, 136, ...

- PCA特徵降維

104, 83, 74, 122, 119, 80, 250, 78, 126, 250, 57, 123, 42, 235, 181, 39, 114, 168, 166, 136, ...

16863個



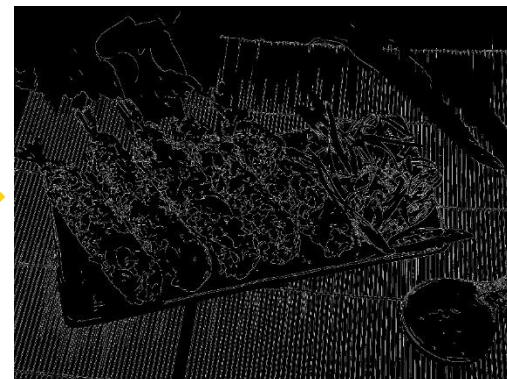
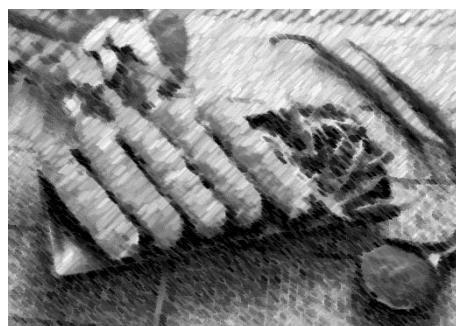
, 57, 123, 42, 235, 181, 39, 114, 168, 166, 136, ...

32個

圖片特徵擷取方式(2)

圖片轉換為128x128像素的灰階影像。

- CANNY取得物體輪廓



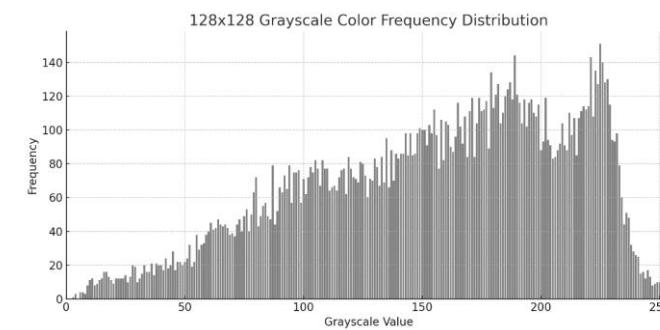
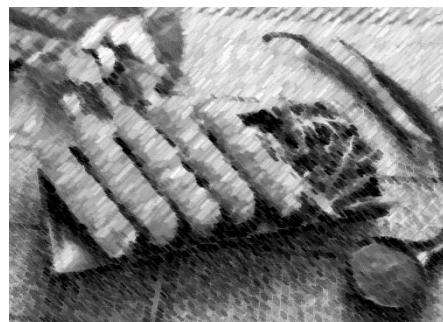
104, 83, 74, 122, 119, 80, 250, 78,
216, 244, 166, 249, 59, 160, 182, 1,
73, 132, 212, 160, 95, 36, 11, 92, ;
243, 91, 183, 254, 19, 168, 158, 232
56, 12, 159, 210, 46, 182, 216, 208
49, 204, 120, 34, 1, 173, 25, 18,
184, 77, 230, 223, 162, 193, 215, 22
140, 119, 206, 22, 3, 107, 143, 16,
122, 199, 192, 71, 33, 209, 43, 242
181, 135, 106, 96, 205, 253, 195, 58,



, 57, 123, 42, 235, 181, 39, 114, 168, 166, 136, ...

- 色彩頻率

計算圖像的灰階色彩頻率分佈，輸入255個特徵的一維向量。

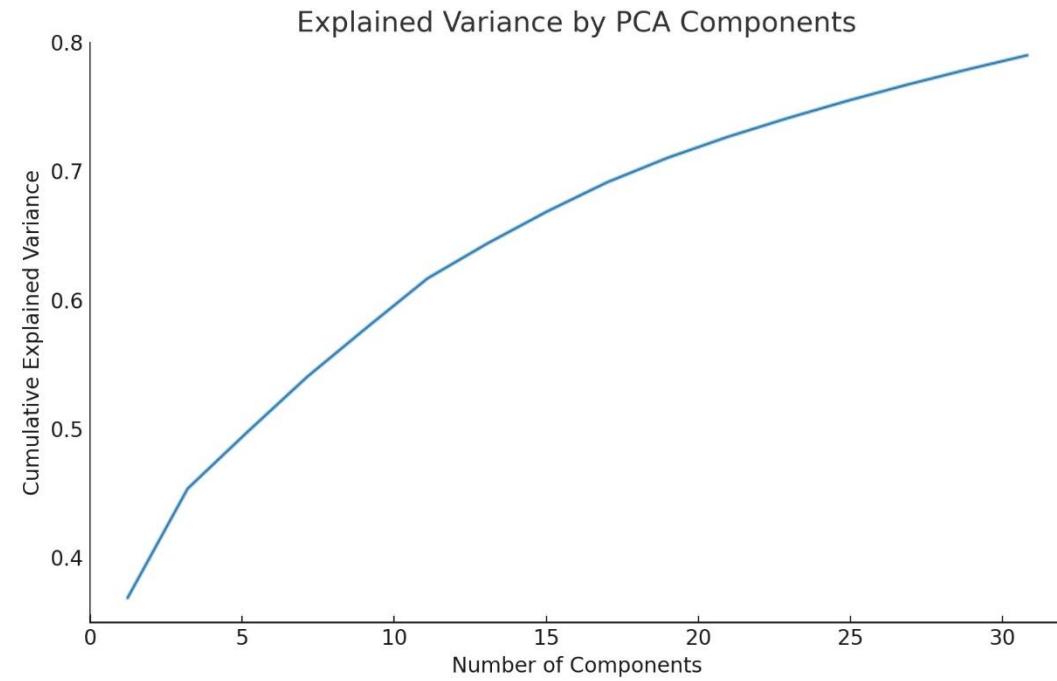


, 57, 123, 42, 235, 181, 39, 114, 168, 166, 136, ...

串聯兩者結果後PCA(32,)矩陣

PCA方差解釋率

降維成32維度後，對於原本的特徵約有80%解釋力。

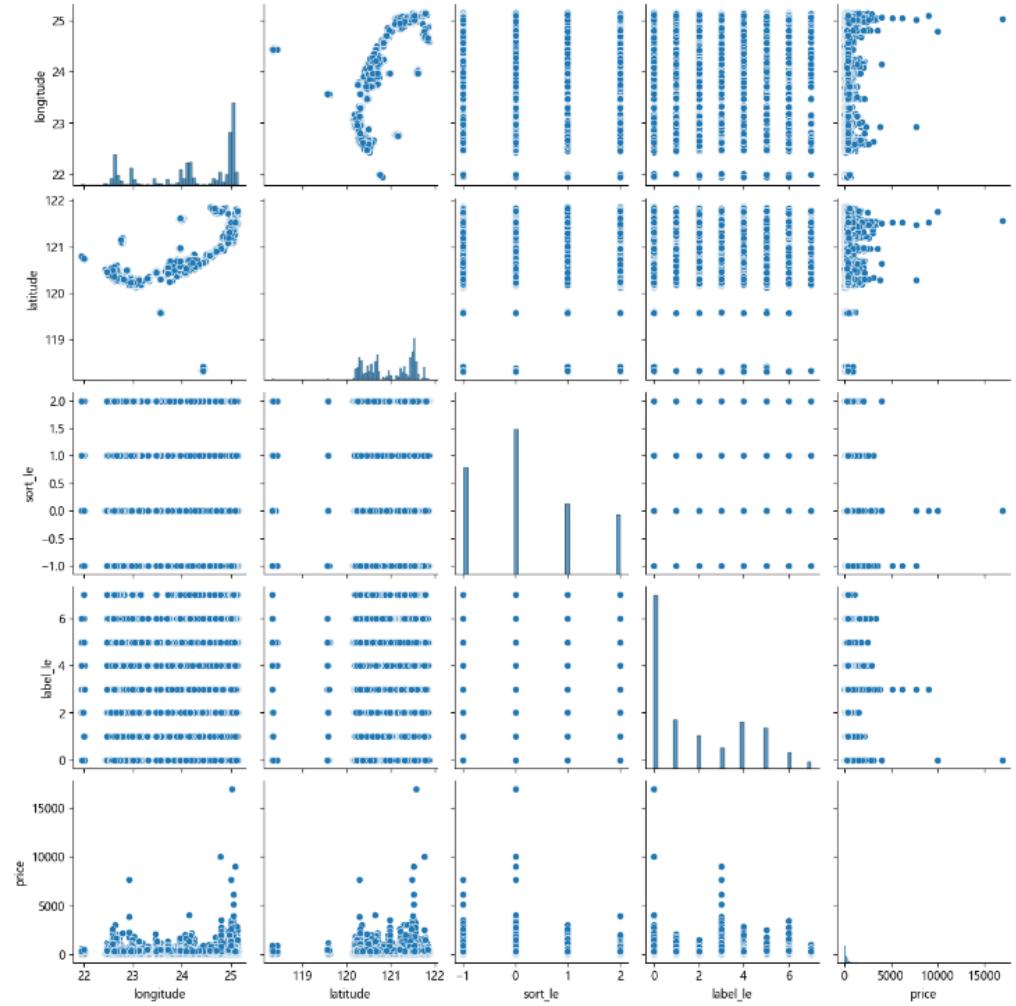
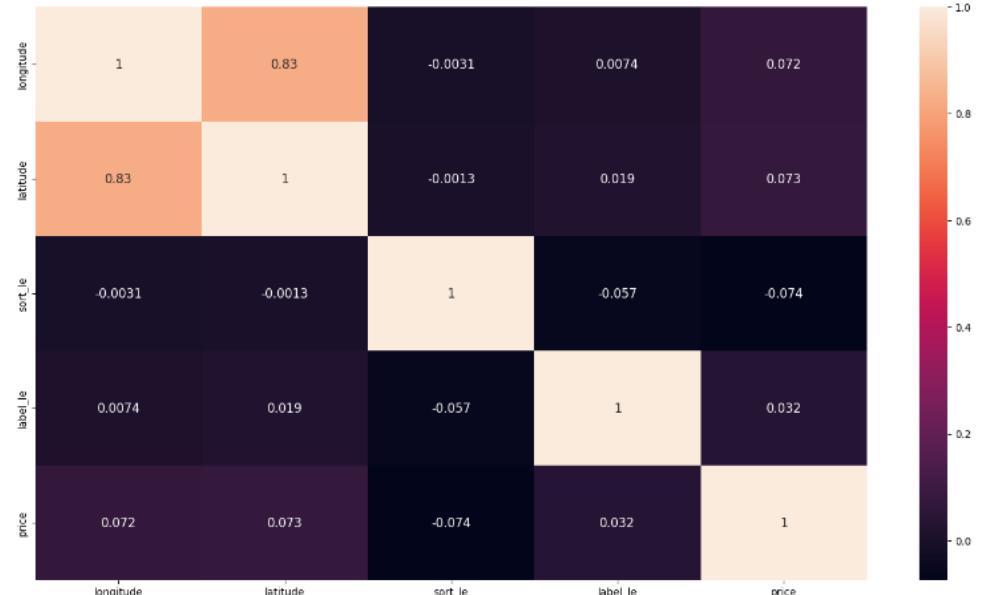


特徵間關係分析

散點圖與熱力圖

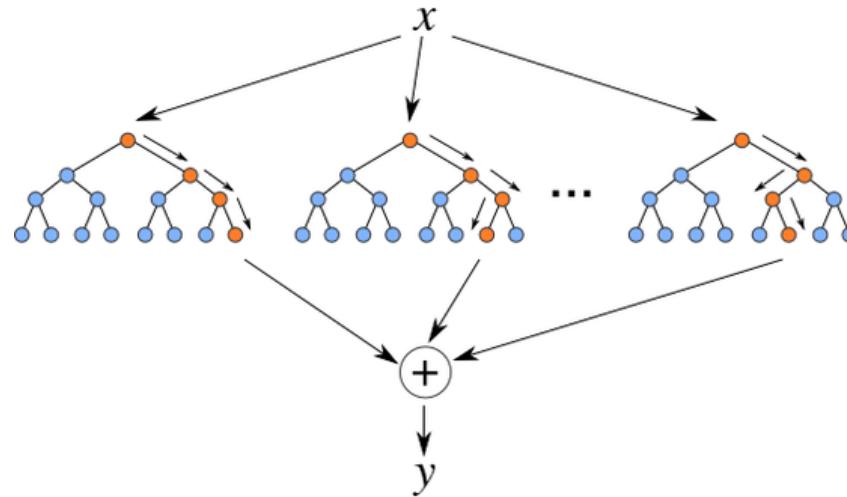
主講者：魯又齊

- 各X與y之間並沒有明顯的關聯性
- 經緯度有相關性

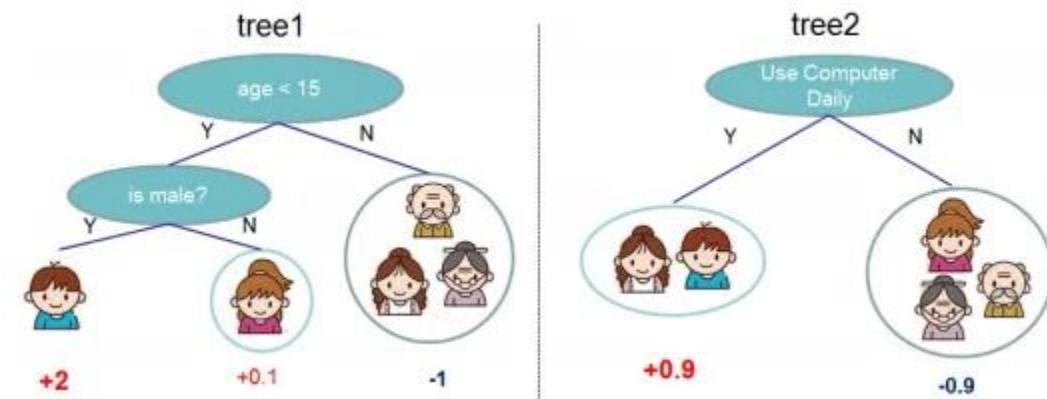


機器學習模型

隨機森林



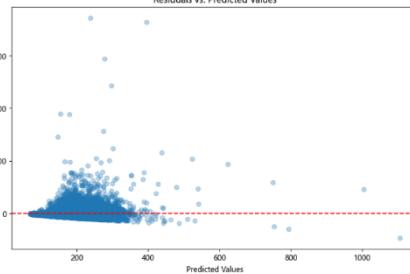
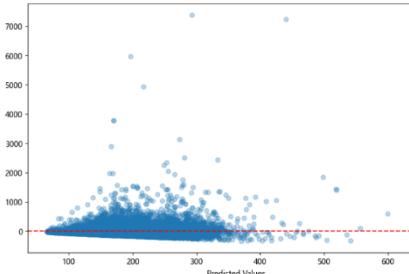
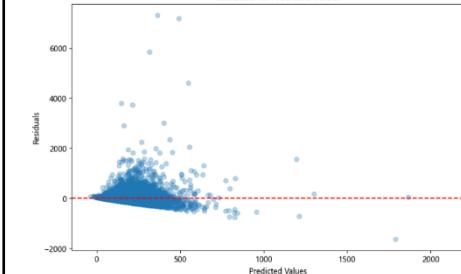
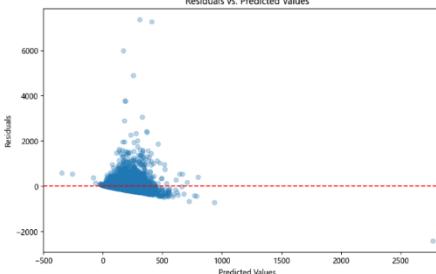
XGBoost



模型效能檢測

回歸(真實價格)

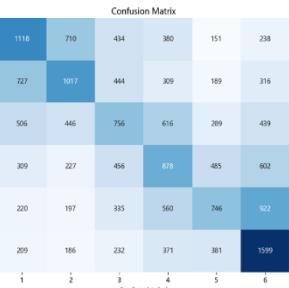
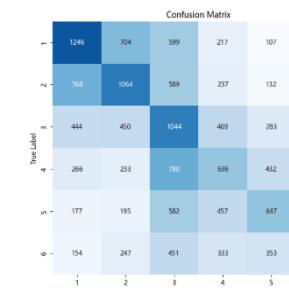
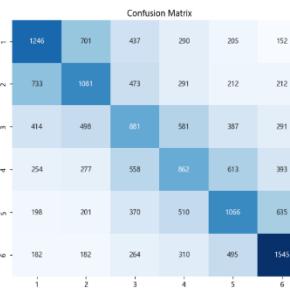
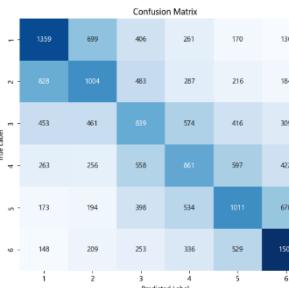
主講者：魯又齊

模型	隨機森林-回歸	隨機森林-CANNY-回歸	XGBOOST回歸	XGBOOST-CANNY-回歸
測試集R ²	0.160	0.166	0.169	0.162
驗證集R ²	0.112	0.126	0.125	0.122
預測殘差分布				

模型效能檢測

分類(價格級距)

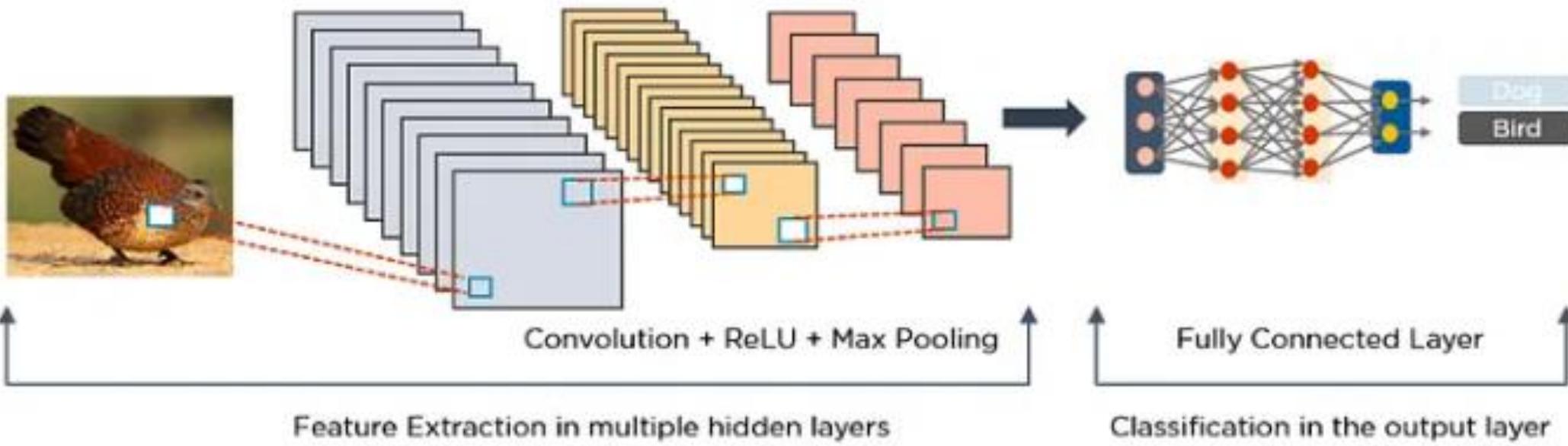
主講者：魯又齊

模型	隨機森林	隨機森林-CANNY	XGBOOST	XGBOOST-CANNY																																																																																																																																																																																																				
測試集準確性	0.338	0.338	0.357	0.351																																																																																																																																																																																																				
驗證集準確性	0.339	0.337	0.371	0.365																																																																																																																																																																																																				
F1 Score	0.334	0.333	0.369	0.363																																																																																																																																																																																																				
混淆矩陣	 <table border="1"> <caption>Confusion Matrix</caption> <thead> <tr> <th>Predicted Label</th> <th>1</th> <th>2</th> <th>3</th> <th>4</th> <th>5</th> <th>6</th> </tr> </thead> <tbody> <tr> <th>True Label 1</th> <td>1118</td> <td>710</td> <td>434</td> <td>360</td> <td>151</td> <td>238</td> </tr> <tr> <th>True Label 2</th> <td>727</td> <td>1017</td> <td>444</td> <td>309</td> <td>189</td> <td>316</td> </tr> <tr> <th>True Label 3</th> <td>506</td> <td>446</td> <td>756</td> <td>616</td> <td>289</td> <td>439</td> </tr> <tr> <th>True Label 4</th> <td>309</td> <td>227</td> <td>456</td> <td>878</td> <td>485</td> <td>902</td> </tr> <tr> <th>True Label 5</th> <td>220</td> <td>197</td> <td>335</td> <td>560</td> <td>746</td> <td>922</td> </tr> <tr> <th>True Label 6</th> <td>209</td> <td>186</td> <td>232</td> <td>371</td> <td>381</td> <td>1599</td> </tr> </tbody> </table>	Predicted Label	1	2	3	4	5	6	True Label 1	1118	710	434	360	151	238	True Label 2	727	1017	444	309	189	316	True Label 3	506	446	756	616	289	439	True Label 4	309	227	456	878	485	902	True Label 5	220	197	335	560	746	922	True Label 6	209	186	232	371	381	1599	 <table border="1"> <caption>Confusion Matrix</caption> <thead> <tr> <th>Predicted Label</th> <th>1</th> <th>2</th> <th>3</th> <th>4</th> <th>5</th> <th>6</th> </tr> </thead> <tbody> <tr> <th>True Label 1</th> <td>1246</td> <td>704</td> <td>599</td> <td>217</td> <td>107</td> <td>158</td> </tr> <tr> <th>True Label 2</th> <td>708</td> <td>1064</td> <td>589</td> <td>237</td> <td>132</td> <td>212</td> </tr> <tr> <th>True Label 3</th> <td>444</td> <td>450</td> <td>1044</td> <td>469</td> <td>283</td> <td>362</td> </tr> <tr> <th>True Label 4</th> <td>269</td> <td>233</td> <td>790</td> <td>636</td> <td>432</td> <td>610</td> </tr> <tr> <th>True Label 5</th> <td>177</td> <td>195</td> <td>582</td> <td>457</td> <td>647</td> <td>922</td> </tr> <tr> <th>True Label 6</th> <td>154</td> <td>247</td> <td>451</td> <td>333</td> <td>353</td> <td>1440</td> </tr> </tbody> </table>	Predicted Label	1	2	3	4	5	6	True Label 1	1246	704	599	217	107	158	True Label 2	708	1064	589	237	132	212	True Label 3	444	450	1044	469	283	362	True Label 4	269	233	790	636	432	610	True Label 5	177	195	582	457	647	922	True Label 6	154	247	451	333	353	1440	 <table border="1"> <caption>Confusion Matrix</caption> <thead> <tr> <th>Predicted Label</th> <th>1</th> <th>2</th> <th>3</th> <th>4</th> <th>5</th> <th>6</th> </tr> </thead> <tbody> <tr> <th>True Label 1</th> <td>1246</td> <td>701</td> <td>437</td> <td>290</td> <td>285</td> <td>152</td> </tr> <tr> <th>True Label 2</th> <td>733</td> <td>1081</td> <td>473</td> <td>291</td> <td>212</td> <td>212</td> </tr> <tr> <th>True Label 3</th> <td>414</td> <td>496</td> <td>811</td> <td>581</td> <td>387</td> <td>291</td> </tr> <tr> <th>True Label 4</th> <td>254</td> <td>277</td> <td>558</td> <td>862</td> <td>613</td> <td>393</td> </tr> <tr> <th>True Label 5</th> <td>198</td> <td>201</td> <td>370</td> <td>510</td> <td>1056</td> <td>615</td> </tr> <tr> <th>True Label 6</th> <td>182</td> <td>182</td> <td>264</td> <td>310</td> <td>405</td> <td>1545</td> </tr> </tbody> </table>	Predicted Label	1	2	3	4	5	6	True Label 1	1246	701	437	290	285	152	True Label 2	733	1081	473	291	212	212	True Label 3	414	496	811	581	387	291	True Label 4	254	277	558	862	613	393	True Label 5	198	201	370	510	1056	615	True Label 6	182	182	264	310	405	1545	 <table border="1"> <caption>Confusion Matrix</caption> <thead> <tr> <th>Predicted Label</th> <th>1</th> <th>2</th> <th>3</th> <th>4</th> <th>5</th> <th>6</th> </tr> </thead> <tbody> <tr> <th>True Label 1</th> <td>1359</td> <td>699</td> <td>406</td> <td>261</td> <td>170</td> <td>136</td> </tr> <tr> <th>True Label 2</th> <td>629</td> <td>1004</td> <td>483</td> <td>287</td> <td>216</td> <td>184</td> </tr> <tr> <th>True Label 3</th> <td>453</td> <td>461</td> <td>619</td> <td>574</td> <td>416</td> <td>309</td> </tr> <tr> <th>True Label 4</th> <td>263</td> <td>256</td> <td>558</td> <td>611</td> <td>597</td> <td>422</td> </tr> <tr> <th>True Label 5</th> <td>173</td> <td>194</td> <td>398</td> <td>534</td> <td>1011</td> <td>670</td> </tr> <tr> <th>True Label 6</th> <td>148</td> <td>209</td> <td>253</td> <td>336</td> <td>529</td> <td>1503</td> </tr> </tbody> </table>	Predicted Label	1	2	3	4	5	6	True Label 1	1359	699	406	261	170	136	True Label 2	629	1004	483	287	216	184	True Label 3	453	461	619	574	416	309	True Label 4	263	256	558	611	597	422	True Label 5	173	194	398	534	1011	670	True Label 6	148	209	253	336	529	1503
Predicted Label	1	2	3	4	5	6																																																																																																																																																																																																		
True Label 1	1118	710	434	360	151	238																																																																																																																																																																																																		
True Label 2	727	1017	444	309	189	316																																																																																																																																																																																																		
True Label 3	506	446	756	616	289	439																																																																																																																																																																																																		
True Label 4	309	227	456	878	485	902																																																																																																																																																																																																		
True Label 5	220	197	335	560	746	922																																																																																																																																																																																																		
True Label 6	209	186	232	371	381	1599																																																																																																																																																																																																		
Predicted Label	1	2	3	4	5	6																																																																																																																																																																																																		
True Label 1	1246	704	599	217	107	158																																																																																																																																																																																																		
True Label 2	708	1064	589	237	132	212																																																																																																																																																																																																		
True Label 3	444	450	1044	469	283	362																																																																																																																																																																																																		
True Label 4	269	233	790	636	432	610																																																																																																																																																																																																		
True Label 5	177	195	582	457	647	922																																																																																																																																																																																																		
True Label 6	154	247	451	333	353	1440																																																																																																																																																																																																		
Predicted Label	1	2	3	4	5	6																																																																																																																																																																																																		
True Label 1	1246	701	437	290	285	152																																																																																																																																																																																																		
True Label 2	733	1081	473	291	212	212																																																																																																																																																																																																		
True Label 3	414	496	811	581	387	291																																																																																																																																																																																																		
True Label 4	254	277	558	862	613	393																																																																																																																																																																																																		
True Label 5	198	201	370	510	1056	615																																																																																																																																																																																																		
True Label 6	182	182	264	310	405	1545																																																																																																																																																																																																		
Predicted Label	1	2	3	4	5	6																																																																																																																																																																																																		
True Label 1	1359	699	406	261	170	136																																																																																																																																																																																																		
True Label 2	629	1004	483	287	216	184																																																																																																																																																																																																		
True Label 3	453	461	619	574	416	309																																																																																																																																																																																																		
True Label 4	263	256	558	611	597	422																																																																																																																																																																																																		
True Label 5	173	194	398	534	1011	670																																																																																																																																																																																																		
True Label 6	148	209	253	336	529	1503																																																																																																																																																																																																		

深度學習

卷積神經網路

主講者：魯又齊

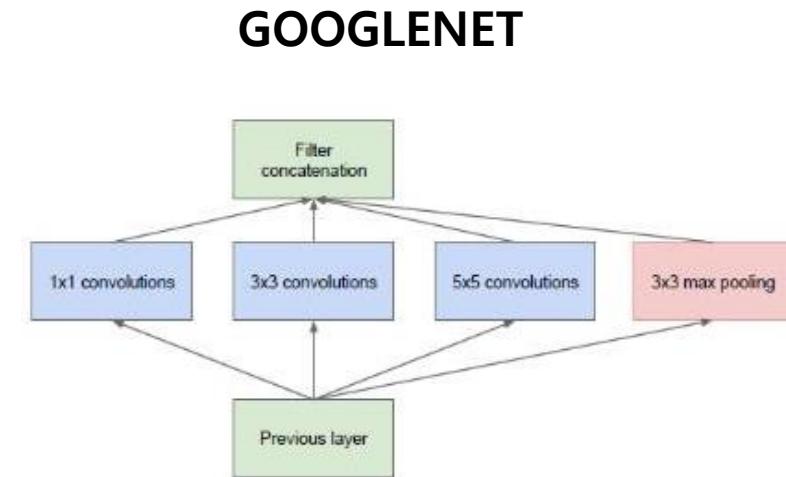
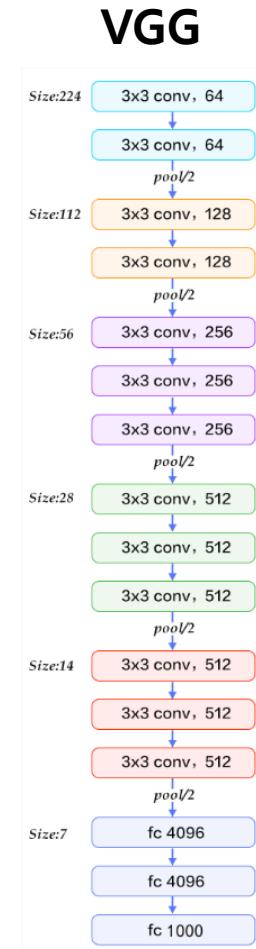
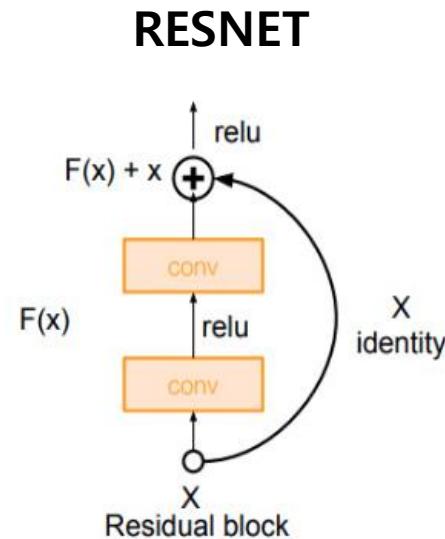


深度學習

預訓練模型

主講者：魯又齊

模型	描述
RESNET	使用「殘差學習」來解決深度網絡中的退化問題
VGG	使用連續的3X3卷積層和2X2池化層
GOOGLENET	不同大小的過濾器捕捉不同視野尺度的特徵

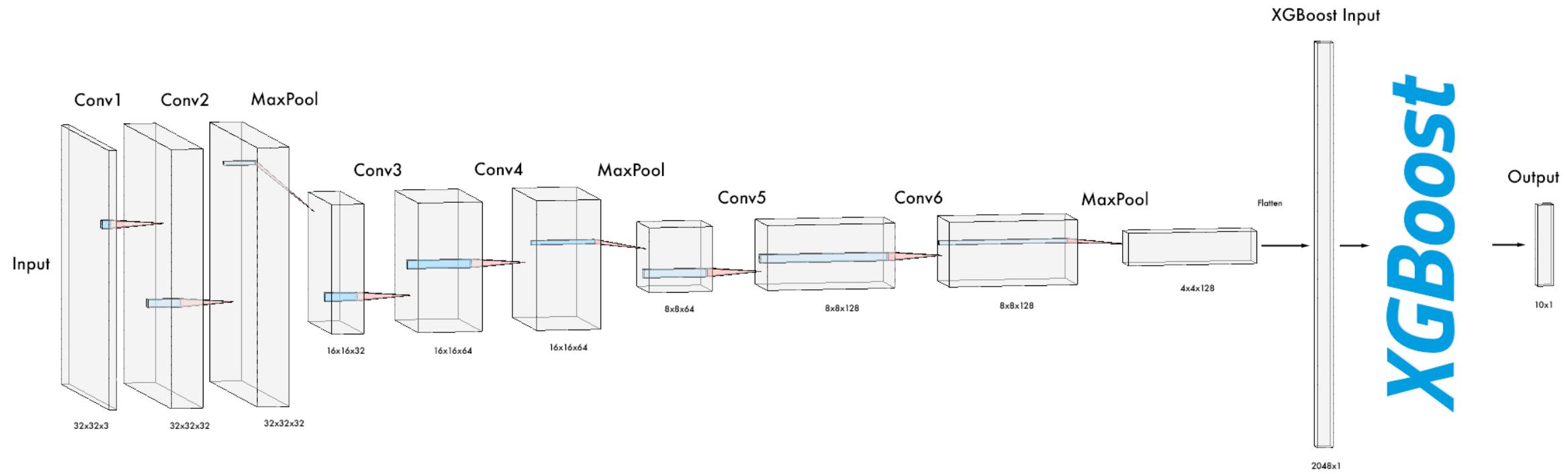


遷移學習

RESNET50/ XGBOOST

主講者：魯又齊

RESNET50



遷移學習

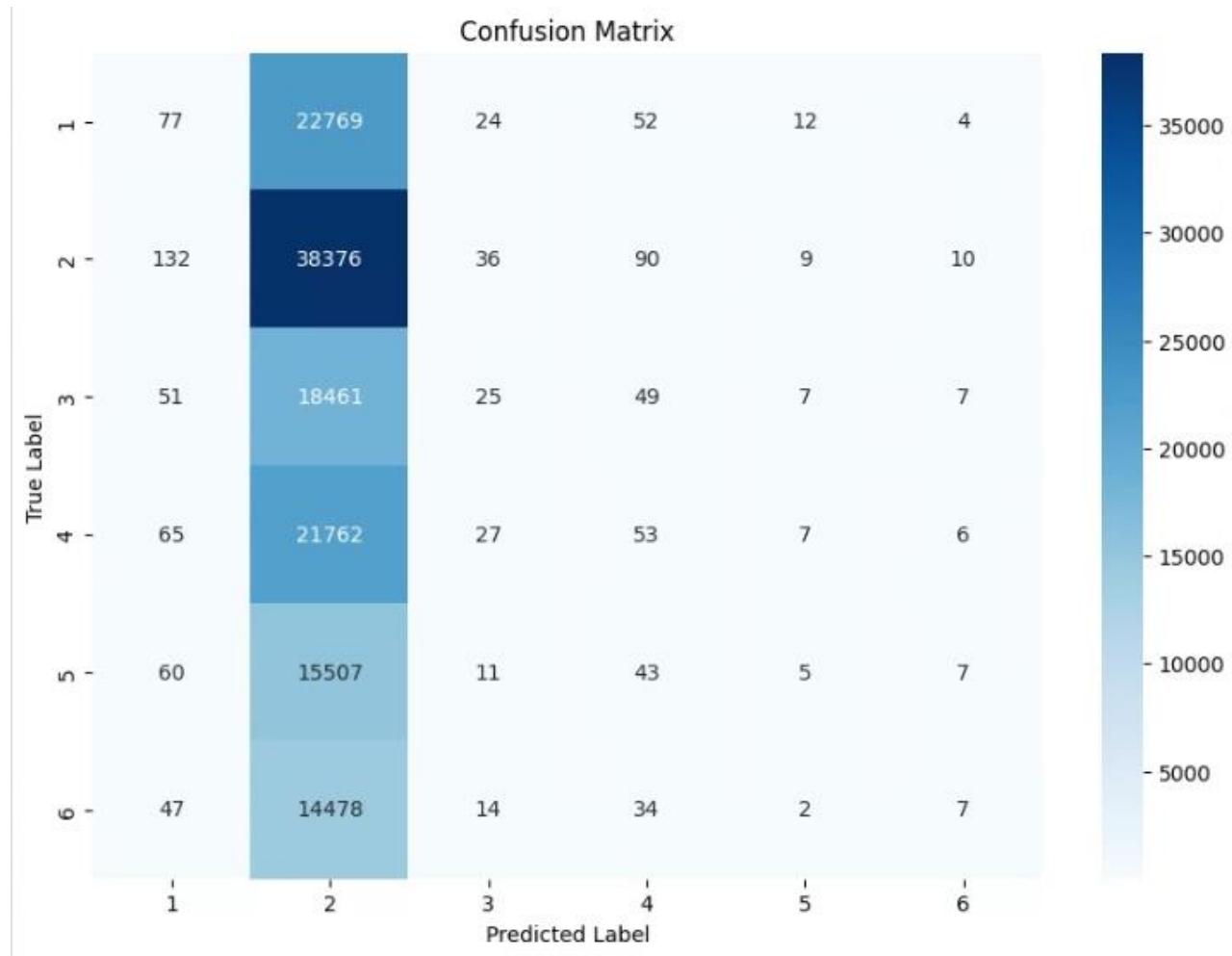
模型評估

主講者：魯又齊

指標

測試Accuracy	0.35
驗證Accuracy	0.29
F1 Score	0.13

混淆矩陣



數據增強

- 訓練圖片尺寸統一



- 較少的類水平及垂直翻轉



- 每個BATCH隨機圖片亮度，飽和度

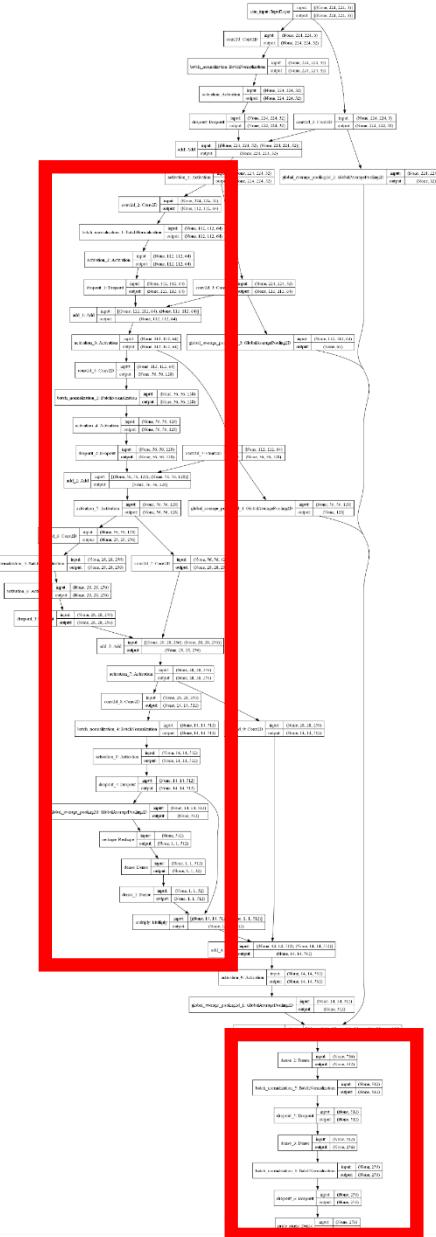


CNN建置

基本層

主講者：魯又齊

層	功能	規劃原因
卷積層	提取圖像特徵	堆疊多個卷積層，模型可以學習到更深的特徵
池化層	降低特徵維度	減少計算量，提高特徵的魯棒性
全連接層	決定最終輸出	學習特徵間的高階關係
正則化層Dropout	隨機關閉一部分神經元	增強泛化能力
正則化層Batch Normalization	保持梯度穩定	減少內部協方差偏移
L1	使模型稀疏	有助於特徵選擇
L2	懲罰模型權重的大小，避免權重過大	防止過擬合



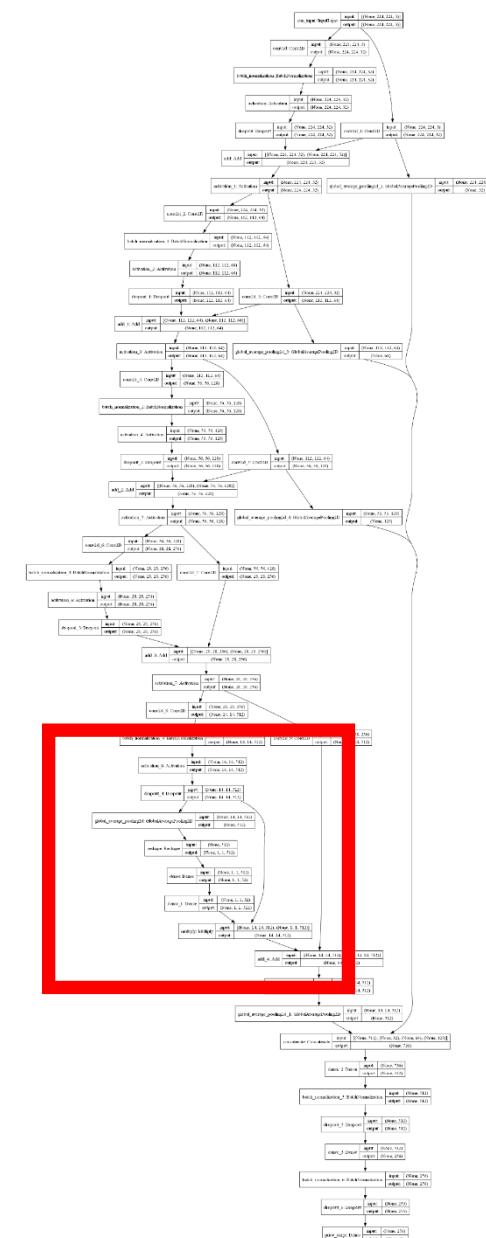
CNN建置

注意力機制

主講者：魯又齊

- SE注意力機制，增加重要特徵權重

功能	技術描述	影響
特徵重構	壓縮 (Squeeze) 激勵 (Excitation)	提升模型對重要特徵的重視程度
特徵選擇	學習每個特徵通道的重要性	強化重要特徵並抑制次要特徵， 增加強化機制
運算效率	使得網絡能夠進行自適應特徵重調	不顯著增加計算負擔

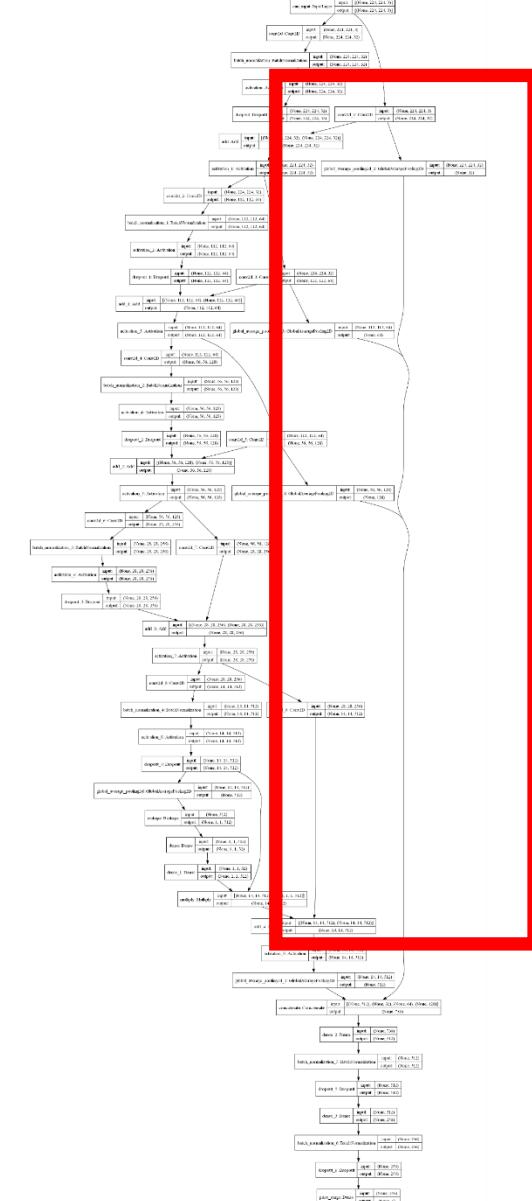


CNN建置

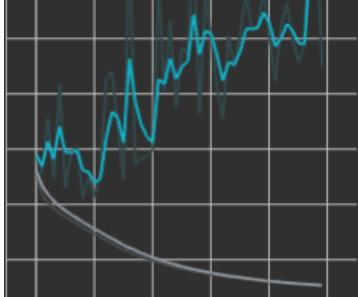
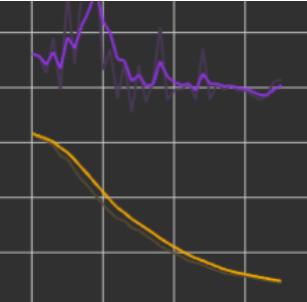
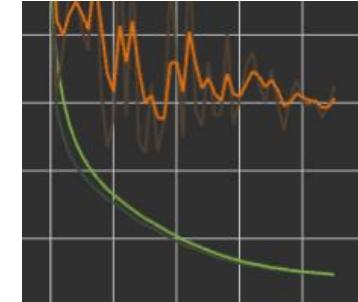
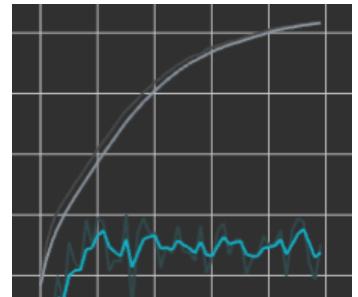
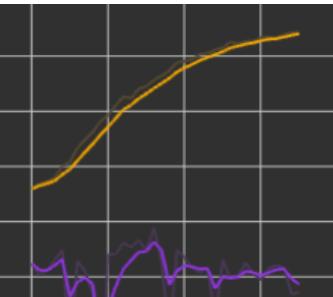
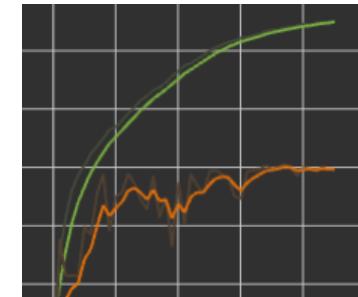
RESNET與DENSENET

主講者：魯又齊

功能	技術描述	影響
特徵融合	融合特徵，強化訊息	增強了模型對重要特徵的識別能力
1x1卷積維度壓縮	調整特徵形狀	提高運算效率
殘差連接	特徵跳躍傳遞	促進梯度流通避免深度網絡的梯度消失
特徵重用	合並前層輸出	提高模型對複雜數據的識別能力



訓練結果

	預訓練 RESNET VGG Inception 224X224X3	自建CNN 224X224X3	自建CNN 324X324X3
BATCH		16	
EPOCHS	50	50	50
LOSS	cross entropy		
優化器	Adam，初始學習率: 0.001，學習率遞減: 每Epoch若發散下降5%		
最大驗證準確度	42~47	52	62
訓練過程 LOSS			
訓練過程 準確度			

雲端串接

後端工程

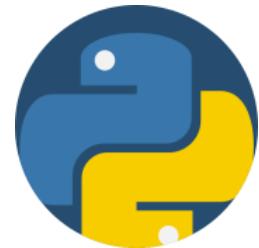


林聖尉
Speaker

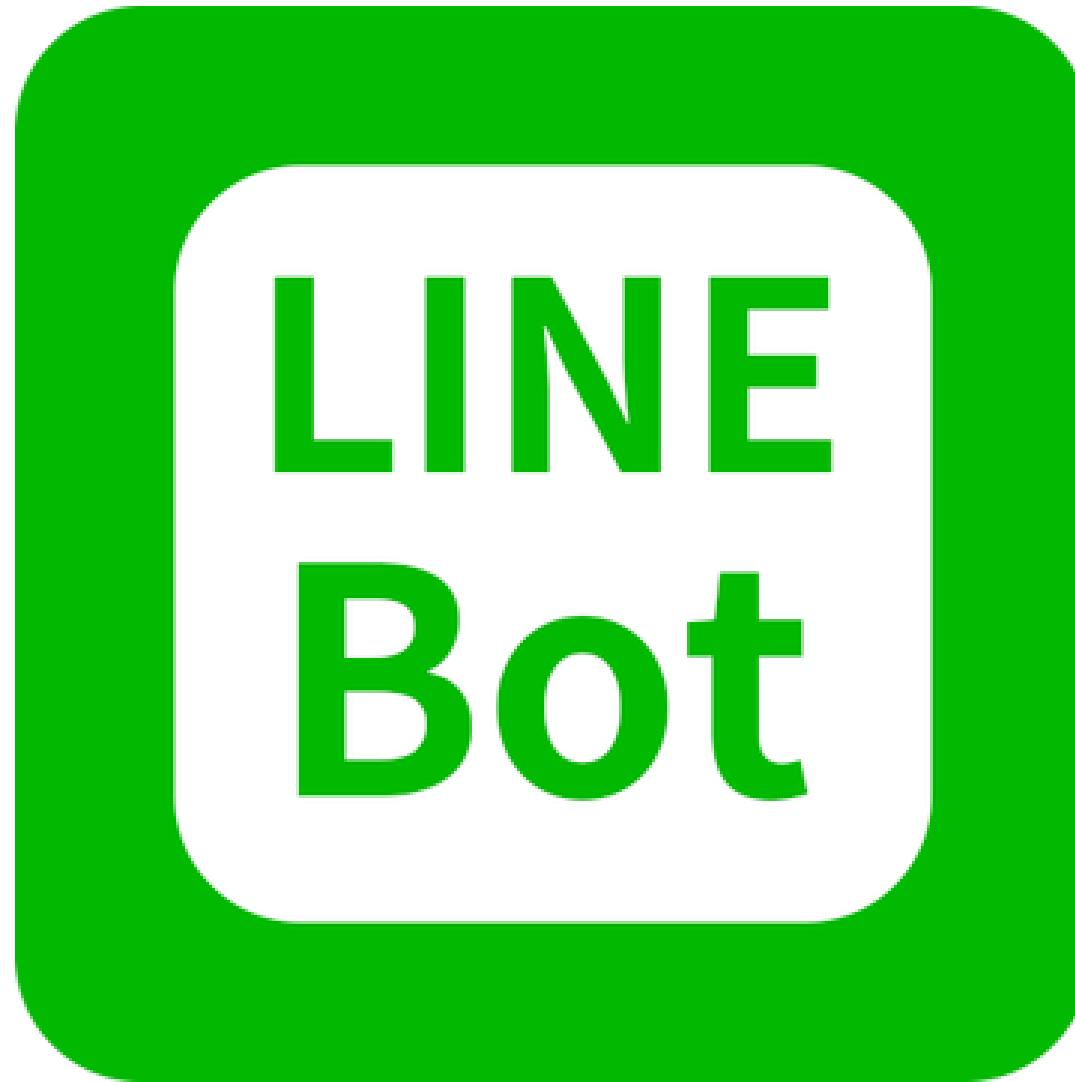
使用工具



Flask



CHATGPT



Azure



Azure Blob
Storage

主講者：林聖尉

系統測試



功能說明



Search

Random

Miee

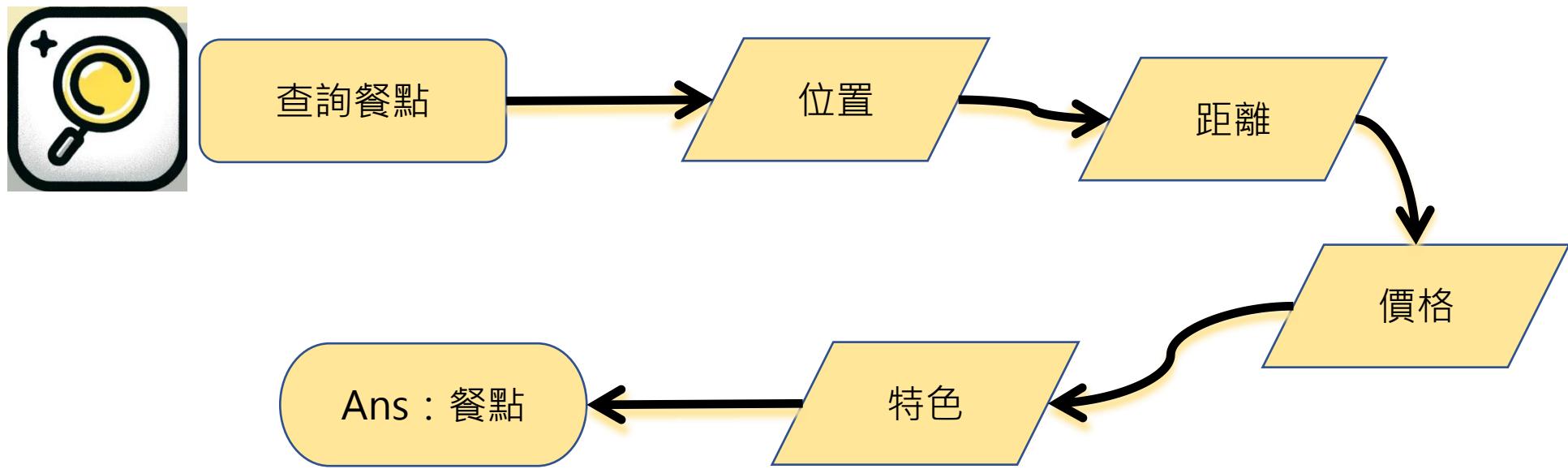
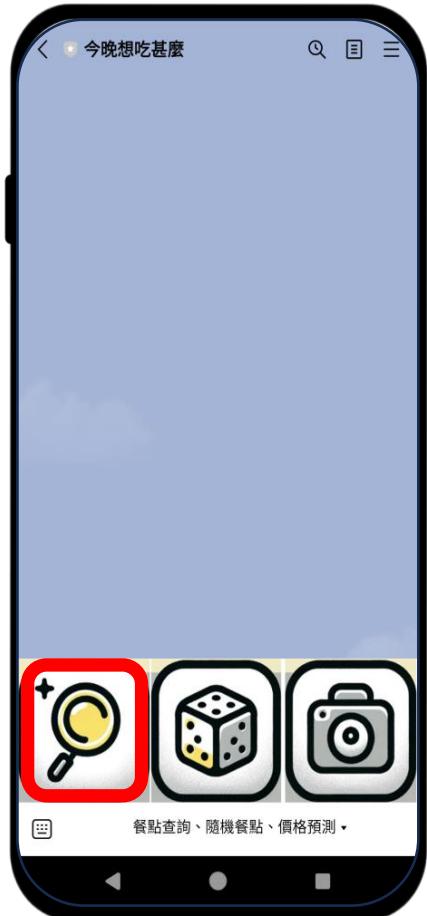
1. 查詢餐點

2. 隨機餐點

3. 價格預測

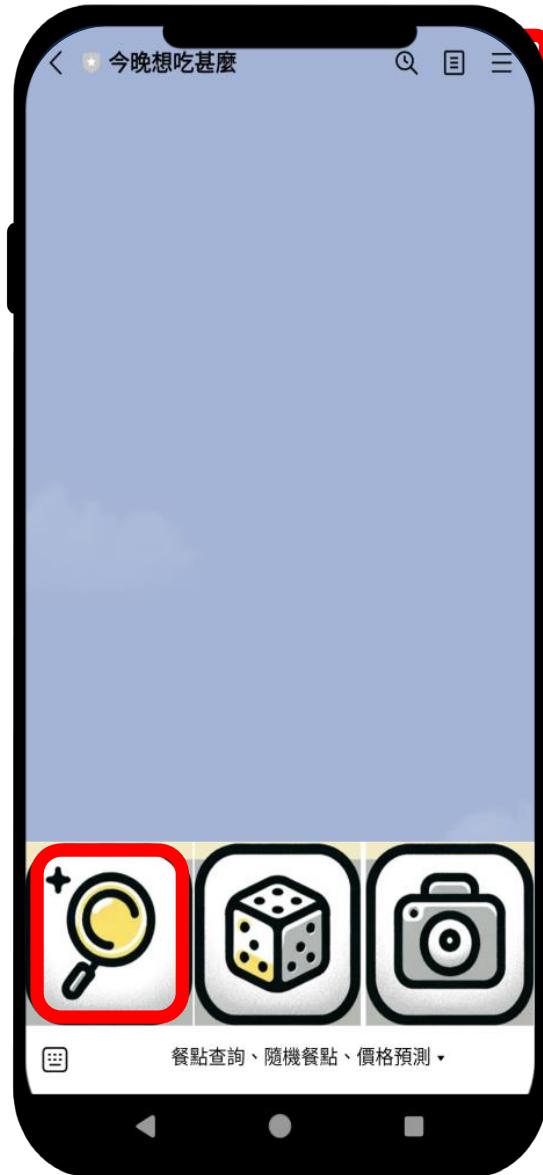
功能說明

1. 查詢餐點



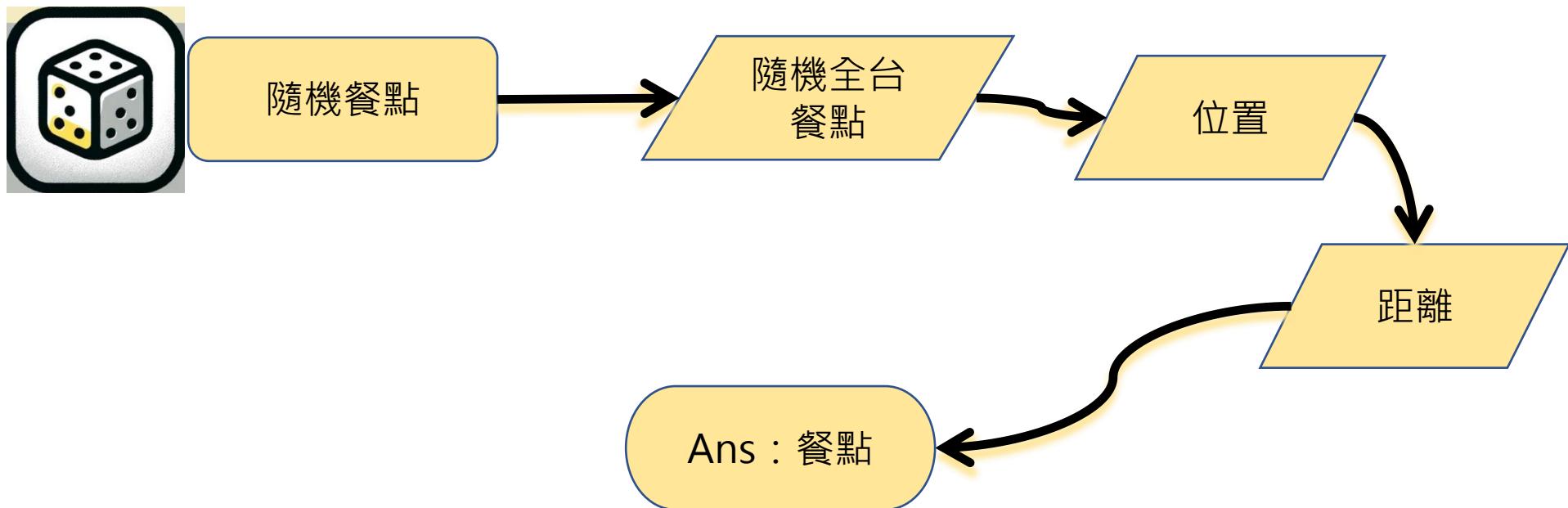
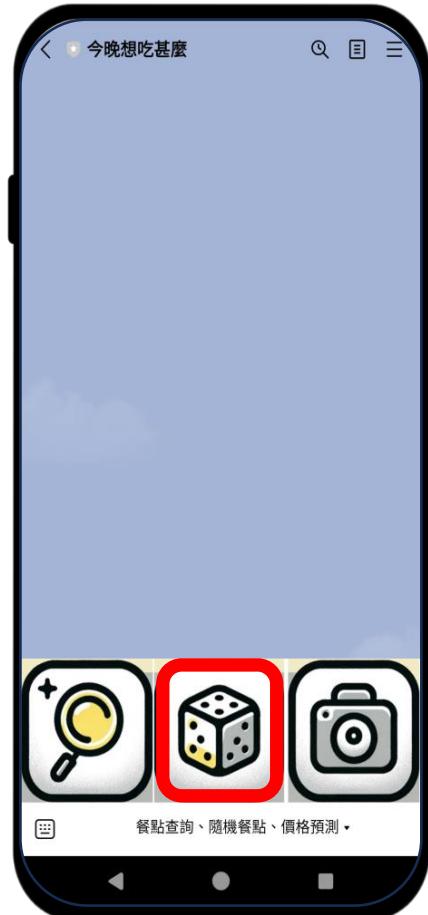
功能說明

1. 查詢餐點 功能示範



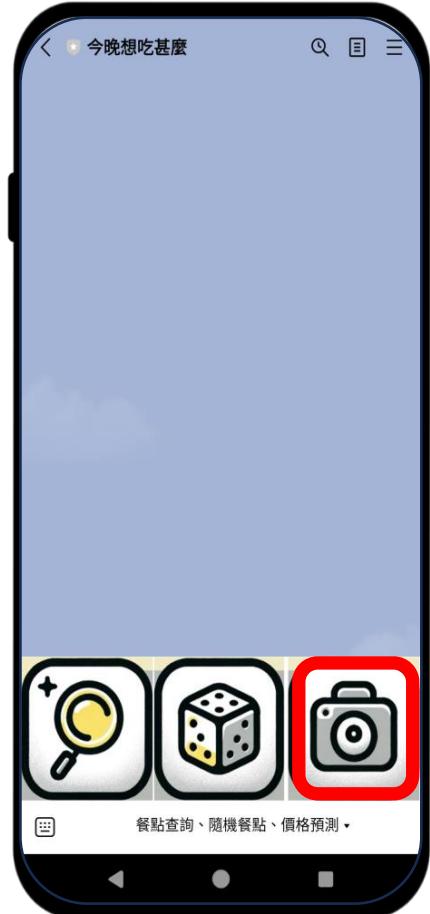
功能說明

2.隨機餐點



功能說明

3. 價格預測



價格預測

上傳圖片

辨識價錢

特色

Ans : 餐點

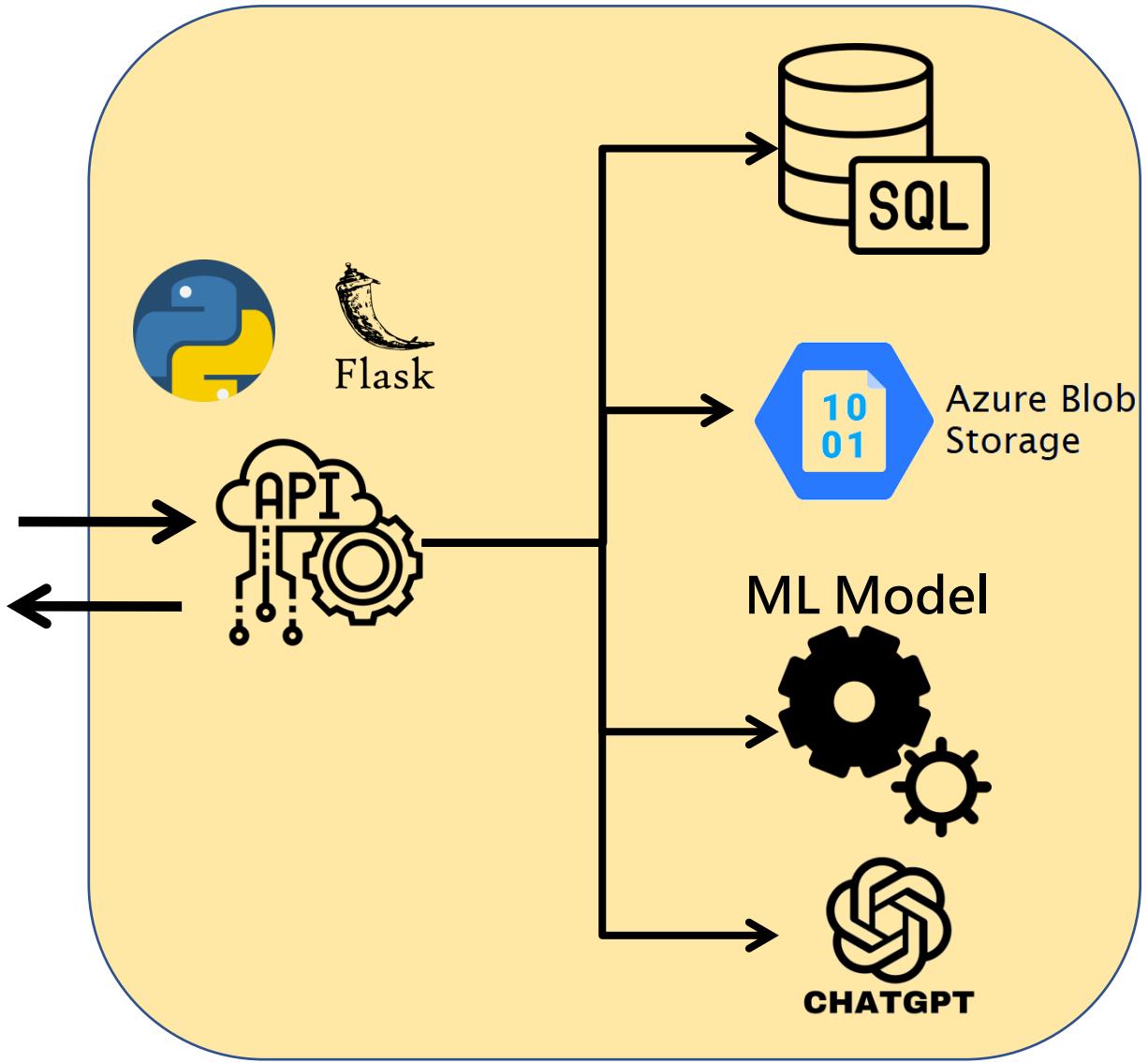
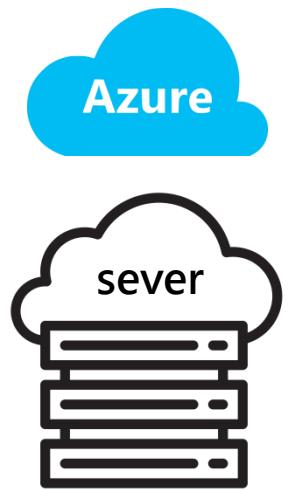
距離

位置

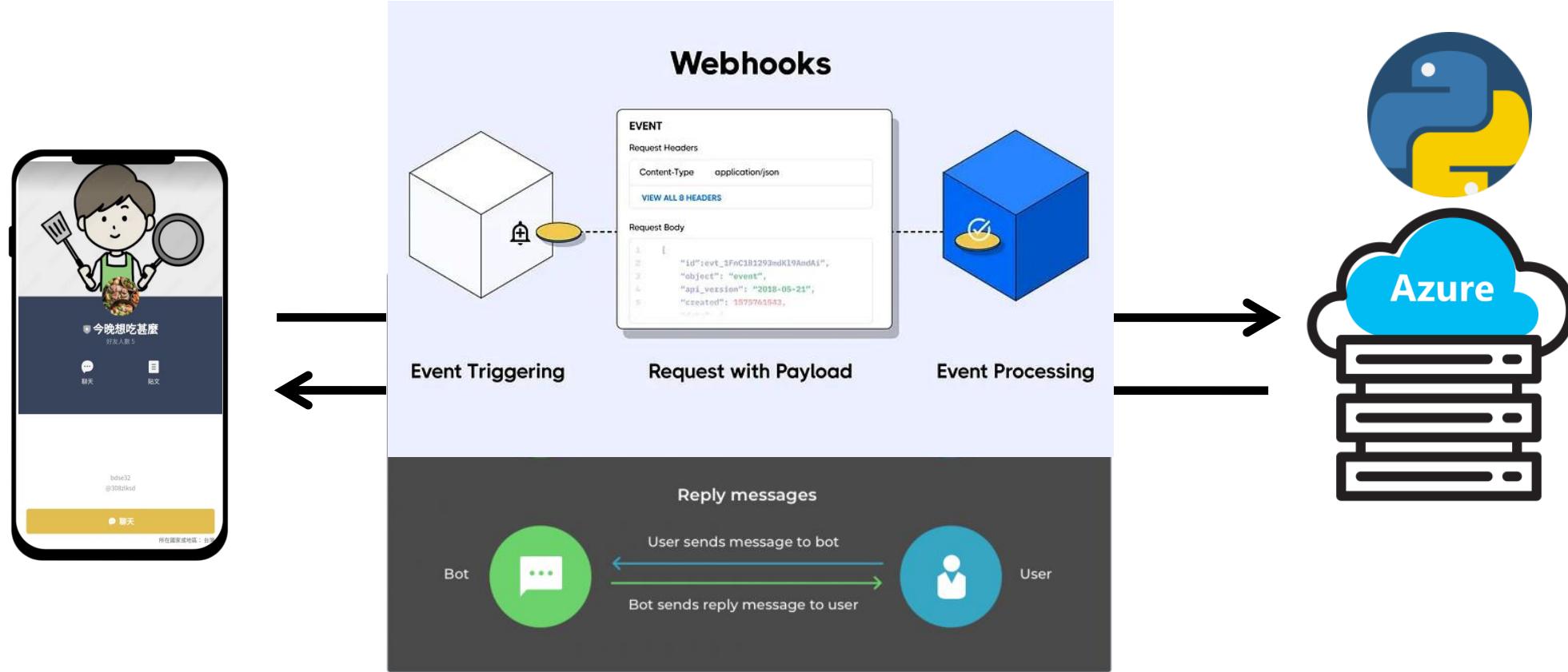
Line Bot架構



Request、Picture
Answer、Picture



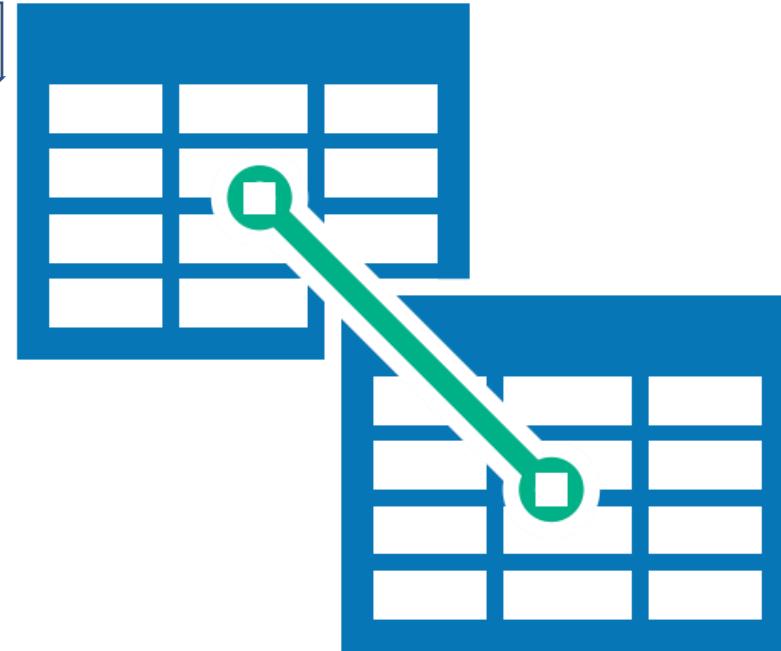
架構說明



架構說明

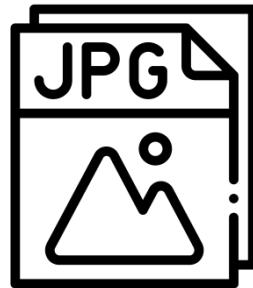


Google commit

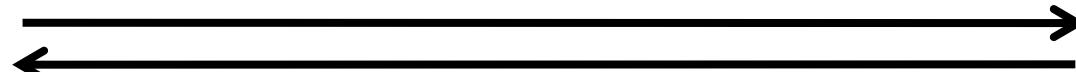


Store data

架構說明



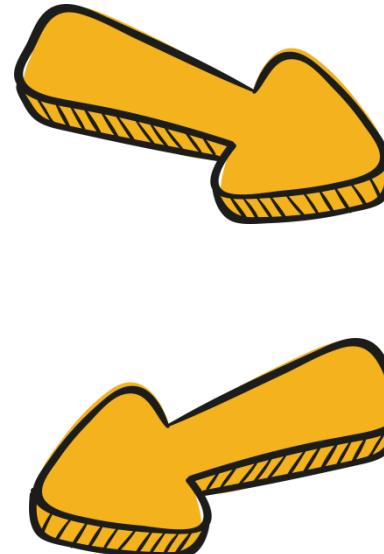
<https://url.jpg>



Azure Blob
Storage



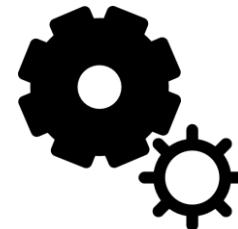
Azure Blob
Storage



架構說明



ML Model

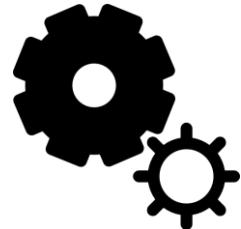


TensorFlow



Input photo

ML Model



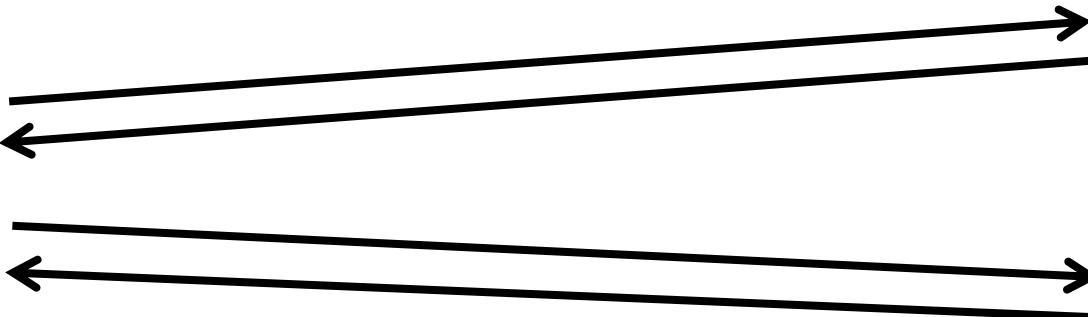
TensorFlow

Answer:300元以上

架構說明



CHATGPT

Azure SQL
Database

Prompt>google commit:

'店名:粵華軒燒臘餐盒',店家評價:店內內用明亮乾淨，三寶便當配菜還可以，但美中不足的是荷包蛋變成炒蛋，主菜三樣肉，叉燒肉和燒鴨蠻夠味的，口感不錯，油雞肉偏硬，140元三寶飯不算便宜，但便當整體感覺和口感都還不錯吃，值得下回再次造訪！環境乾淨，週六晚間客人不多，使用一次用餐具，有湯可盛。弟弟對皮蛋瘦肉粥讚不絕口，烤鴨飯和叉燒公仔麵普普，飯、菜、及肉量少之又少，最扯的是燒肉只有6塊，一個便當居然要140元，下次肯定不會再買。白飯米心沒熟，但是飯居然是爛的！？口感超噁心！我還一次買了四個便當，白飯都沒辦法吃，一個便當賣\$140 這種品管也太爛了吧... 叉燒飯10元 配菜一青菜 一煎蛋 一豆干 少許酸菜 叉燒的量不多 不過叉燒很好吃 不會太乾 軟嫩好吃 醬汁雖不如鳳城 但有自己的特色 米飯還可以牛肉滑蛋滑板118元 牛肉有點紅紅的不知為什麼醬汁 但沒有怪味道 尚可 我點了一份燒肉飯.燒肉切的又少又薄~皮也不脆 ☹咬都咬不動..配菜圖片裡本來應該是要有荷包蛋...後眼~結果竟然是玉米炒蛋 😕難道不知道港式燒肉飯的精神就是荷包蛋嗎！配菜用的青菜又黃又硬..再配上一塊滷的沒人味的豆干..這種貨色136元 😕唉...又貴又難吃...全台北倒數第一了... 烧肉飯130中午賣飯，人不多一個店員負責點單，切肉，等了一會才拿到主菜燒肉不錯，皮脆肉嫩，份量不多跟照片蠻多的配菜普通，豆干偏鹹只有方便，地點優勢 店面明亮乾淨，價位在大安區算是正常。這次點叉燒飯，飯不算多，適合食量小或不想吃太多飯的人，但醬汁頗下飯，所以男性或大食量可能會吃不飽。配菜相當簡單，一塊豆干、荷包蛋、幾根青菜跟蘿蔔乾，頗有香港不澎湃的風格。主食叉燒還不錯，不會太甜、香氣也有，份量一般。我的雞腿為什麼少一半？？中間部位跟頭部被取走，拼湊起來以為顧客看不出來嗎？真的是黑店 1. 點了叉燒飯，售價150元，拿到的時候有點後眼（請看照片）。配菜寒酸，感覺花了125買那支腿肉，25買白飯跟配菜。2. 燒鴨不夠入味，略油。白飯粒粒分明是因為「超級」硬，而且是冷的。是放太久還是煮的時候水加不夠？3. 等候時間約6分鐘（現場有人在等內用的情況下）。4. 吃完整個感覺很空虛，只好再叫外送 135。一顆星給食物，一顆星給乾淨的環境。 1份烤鴨飯130！結果給客人吃爛的菜！而且也沒有附湯！要吃的人自己小心！ 超級不推薦，點烤鴨飯118元，骨頭比肉還多，菜也只有一點點，女生吃完這個便當還是吃不飽。CP值非常低，而且也不美味，可能是剛開幕還是有人潮吧，但真的很不好吃。 超級雷，被新開幕的招牌吸引過來，買了一個三寶飯外帶，結果裡面肉超少，配菜只有一片豆干、一個煎蛋、少量芥蘭菜，根本吃不飽，而且重點的肉也普普通通，根本無法跟廣告單上的米其林聯想在一起，完全是一家靠着廣告標語與充電招牌來吸引顧客的店家... 路過看到烤鴨店乾淨淨，於是外帶了一隻烤鴨999元。分兩大淺盤，沒有提供塑膠袋還要跟消費者索價\$3元。回家吃了發現很普通，鴨肉不肥美鴨皮也不脆、沒什麼油脂的乾柴烤鴨稱不上美味。 說貴的吃不飽的，請有個認知知道自己在哪裡，我是覺得味道很夠，份量也剛剛好，評分的都是要便宜又大碗的南部人？ 1. 地址在大安區復興南路一段293號，捷運大安站走路一

ChatGPT Answer:



CHATGPT

粵華軒燒臘餐盒位於台北市大安區復興南路一段293號，捷運大安站步行一分鐘的優越位置，是一家充滿港式風情的店家。這裡的店內內用明亮乾淨，並提供三寶便當及叉燒飯以及牛肉燴飯等各式餐盒，使用優質食材，口味上既不會太甜也不會太油，而且配料量也能滿足客人的胃口。此外，店內服務人員親切、耐心專業且份量適中，特別是三寶便當140元的價格具有CP值。星期日休餐，適合想在工作日享用午餐或加飯的民眾前往。

商業應用與策略

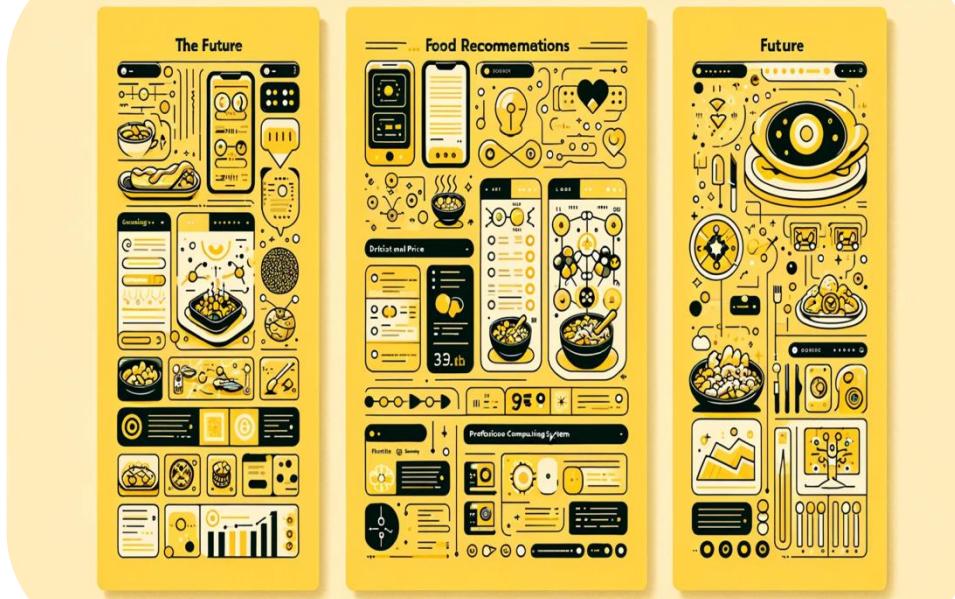


Business Applications & Strategy

□ 潛在商業應用：

- ✓ 為餐廳和食品業者提供市場分析工具。
- ✓ 與訂餐平台應用合作，提供更精準的個性化服務。
- ✓ 提供商家行銷曝光之廣告通路。

未來展望



Future Outlook

1. 更多元的餐點圖片來源，與增加算力與加大訓練圖片提升餐底價格預測功能的能力。
2. 自動化更新資料源。
3. 增加互動平台，不只是透過LINE可以使用到此功能。

Thank you for listening
